



**PREVISION QUANTITATIVE DES
PRECIPITATIONS: ADAPTATION PROBABILISTE
PAR RECHERCHE D'ANALOGUES. Utilisation des
Réanalyses NCEP / NCAR et application aux
précipitations du Sud-Est de la France**

Guillaume Bontron

► **To cite this version:**

Guillaume Bontron. PREVISION QUANTITATIVE DES PRECIPITATIONS: ADAPTATION PROBABILISTE PAR RECHERCHE D'ANALOGUES. Utilisation des Réanalyses NCEP / NCAR et application aux précipitations du Sud-Est de la France. Météorologie. Institut National Polytechnique Grenoble (INPG), 2004. Français. NNT: . tel-01090969

HAL Id: tel-01090969

<https://theses.hal.science/tel-01090969>

Submitted on 5 Dec 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

INSTITUT NATIONAL POLYTECHNIQUE DE GRENOBLE

N° attribué par la bibliothèque

/ / / / / / / / / /

THESE

pour obtenir le grade de

DOCTEUR DE L'INPG

Spécialité : *Océan, Atmosphère, Hydrologie*

préparée au *Laboratoire d'étude des Transferts en Hydrologie et Environnement*
(LTHE, UMR 5564, CNRS-INPG-IRD-UJF)

dans le cadre de l'Ecole Doctorale «*Terre, Univers, Environnement*»

présentée et soutenue publiquement

par

Guillaume BONTRON

le 23 janvier 2004

**PREVISION QUANTITATIVE DES PRECIPITATIONS :
ADAPTATION PROBABILISTE PAR RECHERCHE D'ANALOGUES.**

Utilisation des Réanalyses NCEP / NCAR et application aux précipitations du Sud-Est de la France

JURY

M. M. LANG	Ing. Div., CEMAGREF Lyon	Rapporteur
M. R. ROSSET	Professeur, UPS Toulouse	Rapporteur
M. A. BARDOSSY	Professeur, Universität Stuttgart (D)	Examineur
M. E. BRUN	Dir. de la Recherche, METEOFRANCE Toulouse	Examineur
M. R. GARCON	Ing. expert, EDF-DTG	Examineur
M. F.-X. LE DIMET	Professeur, UJF Grenoble	Examineur
M. Ch. OBLED	Professeur, INP Grenoble	Directeur de thèse

Remerciements

Une fois arrivé au terme de ce travail, une fois que l'étape de la soutenance a été franchie et que la course finale de la rédaction a été digérée, il est temps de se retourner pour enfin apprécier le chemin parcouru. Et ce chemin, je ne l'ai pas effectué seul...

Je tiens tout d'abord à remercier Michel Vauclin et Jean-Dominique Creutin, les deux directeurs successifs du Laboratoire d'étude des Transferts en Hydrologie et Environnement, qui m'ont accueilli au sein de leur laboratoire durant ces 3 années *et quelque* et m'ont ainsi permis de goûter aux joies et aux angoisses de la recherches.

Je souhaite ensuite exprimer toute ma gratitude à mon directeur de thèse, l'inénarrable professeur Charles Obled, qui fut l'enthousiaste instigateur de cette thèse. Toujours disponible (mais après 18 heures, une fois les élèves partis) pour apporter ses avis/conseils/directives scientifiques de haut niveau, mais aussi pour m'encourager dans ces moments de « *petite forme* » que les thésards connaissent épisodiquement. Je garde un souvenir enthousiaste de nos discussions scientifiques et de ses très pédagogiques performances d'acteur dans l'explication des difficultés de la météorologie (ah, la particule d'air chaud...), mais je tiens aussi à saluer ses qualités humaines (qui ont rendu notre coopération si motivante), ainsi que sa parfaite maîtrise de l'art délicat de la digression (ou comment passer du coq à l'âne sans en avoir l'air).

La Division Technique Générale de EDF m'a apporté une aide précieuse, en mettant à ma disposition un grand nombre de données. Je remercie ainsi plus particulièrement Rémy Garçon, Joël Gailhard et Maud Laplanche, qui en plus des données, m'ont accordé du temps pour me conseiller ou partager leurs points de vue.

Abdelatif Djerboua et Olivier Brasseur, bien que post-doctorants sur d'autres projets, ont suivi de très près mon travail et m'ont souvent apporté leur aide précieuse, leur soutien attentif et leurs encouragements stimulants. Qu'ils en soient ici remerciés.

Il me semble important d'associer à ce travail les stagiaires avec qui j'ai eu la chance de pouvoir travailler : Kamal, Jérôme et Ferdinand. Leur travail m'a été très utile et les échanges que l'on a eut m'ont beaucoup apporté. J'espère qu'eux aussi ont trouvé cette collaboration bénéfique et je leur souhaite bon vent pour la suite.

Le LTHE est un peu une grande famille sur laquelle on peut toujours compter. Je tiens ainsi à remercier Sylvianne, Odette et Agnès, les trois fées du secrétariat qui font des miracles malgré les

turpitudes des contraintes administratives, ainsi que Bruno et Hervé, qui m'ont sauvé plus d'une fois de problèmes informatiques insurmontables. Je ne me risquerais pas à citer nommément tous les membres du LTHE avec qui j'ai partagé tous ces moments qui font la saveur de la vie d'un laboratoire (réflexions scientifiques, discussions philosophiques, anecdotes croustillantes), de peur d'en oublier tellement ils sont nombreux. Je ne peux cependant m'empêcher d'évoquer Monsieur le Professeur Bois, à qui je présente à nouveau toutes mes excuses pour les prévisions météorologiques désastreuses faites à l'occasion du mariage de son fils...

Dans la population d'un laboratoire, on compte aussi nombre de thésards. On tisse entre nous des liens importants, liés au partage de moments de doute (souvent), de moments de réussite (si si, ça arrive!), de philosophie de comptoir autour d'un café (plusieurs fois par jours), de sandwiches du R.U. (il faut bien se nourrir ...) et de soirées plus ou moins arrosées. Merci donc à Alexis, Gaël, Fabien, Isabella, Willy, Helena, José, Adam, Marine, Romain, Maud, Noémie, Guillaume, David, Laetitia, Eddy, Théo, Mathieu qui ont fortement contribué à rendre ces 3 années et quelque si enrichissantes.

Enfin, il me semble que si je suis arrivé à bon port, je le dois en grande partie à la présence et au soutien de ma compagne Gwénaëlle. Je la remercie donc de tout mon cœur, et notamment pour avoir eu la patience d'attendre que je rende mon manuscrit avant de mettre au monde Elouan et me faire papa.

Résumé

Ce mémoire étudie l'application du principe d'*analogie* à la prévision et à la simulation des précipitations. À partir de la caractérisation d'une situation météorologique cible, cette approche consiste à rechercher, dans une archive météorologique, les situations passées similaires. Les précipitations observées lors de ces situations analogues permettent ensuite de donner une estimation probabiliste des précipitations attendues pour la situation cible.

Après avoir défini les critères et les scores qui nous permettent de comparer deux prévisions probabilistes, nous avons cherché à identifier les algorithmes d'analogie les plus informatifs vis à vis des précipitations. Pour ce faire, nous nous sommes appuyé sur les réanalyses NCEP/NCAR. La richesse de cette archive nous a permis de définir des algorithmes d'analogie qui intègrent à la fois (i) l'information à grande échelle contenue dans la circulation synoptique, et (ii) l'information à échelle plus locale contenue dans les variables d'humidité.

Les différents paramètres de la méthode (choix des variables, domaine à prendre en compte, ...) ont été optimisés lors d'une phase de calibration. La validation des algorithmes obtenus a ensuite permis d'apprécier les performances atteintes en adoptant le point de vue d'un utilisateur, ainsi que d'étudier la sensibilité de ces algorithmes aux caractéristiques de l'archive météorologique utilisée. Enfin, grâce à la mise en place d'une chaîne opérationnelle prototype, nous avons pu tester les performances opérationnelles de l'adaptation par analogie, notamment via l'étude de cas de l'évènement intense survenu les 8 et 9 septembre 2002 dans les Cévennes.

Abstract

This work deals with the application of the analogy principle to rainfall forecast and simulation. Given a target meteorological situation, similar past situations are searched for in a meteorological archive. The precipitation amounts, associated with those analogous situations, allowed us to build a probabilistic estimation of the expected rainfall amount for the target situation.

At first, the criteria and scores we used to compare two different probabilistic forecasts have been defined. Then, we tried to identify the analogy algorithm which best suits the precipitation forecast. Relying on the NCEP/NCAR meteorological reanalyses, we built analogy algorithms which take into account both (i) large-scale information brought by the synoptic circulation and (ii) more local scale information brought by the humidity variables.

The different parameters of the algorithms (choice of variables, spatial domain to consider, ...) have been optimised through a first calibration step. Further, the validation of the calibrated algorithms allowed to appraise the forecast performances from the user point of view, as well as to study the sensibility of these algorithms to the features of the meteorological archives. Lastly, thanks to the implementation of an operational forecasting chain, we have been able to assess the on-line performances of the adaptation by analogues sorting. As a case-study, we focussed on the catastrophic 8th and 9th of September 2002 hydro-meteorological event, which occurred in the Cévennes area (France).

Table des matières

Avant-propos	1
--------------	---

partie I

CONTEXTE & OUTILS

1 L'approche par analogie en prévision météorologique	7
1.1 Besoins et perspectives d'utilisation	8
1.1.1 Les échéances de prévision	8
1.1.2 Le besoin d'adaptation locale	11
1.1.3 L'émission d'une prévision	12
1.2 Les différentes approches	15
1.2.1 Prévisions objectives et interprétations subjectives	15
1.2.2 La modélisation numérique (NWP)	15
1.2.3 Les approches statistiques	17
1.2.4 Les approches mixtes : prévision d'ensemble et techniques d'adaptation	18
1.3 La prévision par analogie synoptique	23
1.3.1 Qu'est-ce qu'une <i>analogie</i> ?	23
1.3.2 Utilisation de l'analogie en prévision	25
1.3.3 Etat de référence (travaux Guilbaud, 1997)	29
1.3.4 Pourquoi une prévision probabiliste?	34
1.3.5 Objectif de la thèse	37
2 Les données utilisées	39
2.1 L'archive météorologique	39
2.1.1 Contraintes et contradictions	39
2.1.2 L'archive opérationnelle EDF	41
2.1.3 Les Réanalyses	42
2.2 L'archive pluviométrique	46
2.2.1 Du cumul des précipitations	46
2.2.2 Les groupements pluviométriques	48
2.2.3 Transformation des précipitations	51
2.3 Format d'archivage	52
3 Qualité et Utilité d'une prévision	55
3.1 A quoi reconnaît-on une <i>bonne</i> prévision?	56
3.1.1 Qualité et utilité : une simple différence sémantique?	56
3.1.2 Généralités méthodologiques	57

3.1.3	Attributs d'une prévision	59
3.1.4	Tout est relatif	64
3.2	Qualité des prévisions déterministes	65
3.2.1	Prévision par catégories	65
3.2.2	Prévision continue	72
3.3	Utilité des prévisions probabilistes	76
3.3.1	Spécificité des prévisions probabilistes	76
3.3.2	Transformation en prévision déterministe	77
3.3.3	Le prêt-à-porter (quelques critères classiques)	79
3.3.4	Le prêt-à-porter avec retouches (utilisation de la Vraisemblance)	83
3.3.5	Le Sur-Mesure (de la <i>Fonction de Coût</i> au <i>Coût Effectif Sociétal</i>)	87
3.4	Choix arbitrairement retenus	100
3.4.1	L'optimisation des algorithmes	101
3.4.2	Le diagnostic des performances finales	102

partie II

OPTIMISATION DE LA MÉTHODE

Introduction à la PARTIE II	105
4 Les géopotentiels comme variable d'analogie principale	107
4.1 Exploration préliminaire	108
4.1.1 Les critères d'analogies	108
4.1.2 Résultats préliminaires	112
4.2 Utilisation d'un unique champ de géopotentiels	114
4.2.1 Performances moyennes sur l'ensemble des groupements	114
4.2.2 Influence du choix du groupement cible	118
4.3 Combinaison de plusieurs champs de géopotentiels	123
4.3.1 Comment prendre en compte plusieurs champs?	123
4.3.2 Les meilleures combinaisons!	125
4.4 Conclusions	129
5 Du choix de la fenêtre d'analogie	131
5.1 Fenêtres d'analogie statiques	131
5.1.1 Représentativité de chaque point de grille	132
5.1.2 Algorithme d'optimisation	136
5.1.3 Fenêtres communes à l'ensemble des groupements	138
5.1.4 Dépendance à la position géographique du groupement	141
5.2 Conditionnement par la situation cible	147
5.2.1 Constat préliminaire	147
5.2.2 Une fenêtre dynamique?	149
5.2.3 Classification par type de temps	155
5.3 Conclusion	165
6 Variables de second ordre	169
6.1 Stratégie adoptée	170
6.1.1 Définition d'une analogie en deux temps	170
6.1.2 Tri préliminaire	171

6.1.3	Définition de variables élaborées	175
6.2	PW% comme variable de second ordre	180
6.2.1	Choix de l'heure de mesure	181
6.2.2	La fenêtre d'analogie de second ordre	182
6.2.3	Choix de $N1$ et $N2$	186
6.2.4	Régionalisation du choix de la variable de second ordre	188
6.2.5	Importance de la qualité de l'analogie de premier ordre	188
6.3	Approche dynamique	190
6.4	Conclusions	193

partie III

ETUDES DIAGNOSTIQUES ET APPLICATIONS

7	Validation & critique des performances	199
7.1	Validation	200
7.1.1	Méthodologie	200
7.1.2	Algorithmes retenus	200
7.1.3	Comparaison des groupements	201
7.1.4	Validation des algorithmes	204
7.2	Etude des performances	205
7.2.1	Pouvoir discriminant en fonction d'un seuil	207
7.2.2	Choix d'un quantile de référence	209
7.2.3	Exemple de prévisions	211
7.3	Correction du biais ?	214
7.4	Conclusions	215
8	Sensibilité à l'archive météorologique	217
8.1	Homogénéité des archives météorologiques	217
8.1.1	Critique de l'homogénéité	218
8.1.2	Relation entre homogénéité et performance	222
8.2	Longueur d'archive nécessaire	225
8.3	Effet de la résolution spatiale	227
8.4	Conclusions	229
9	Utilisation opérationnelle : étude de cas	231
9.1	Les prévisions opérationnelles au LTHE	232
9.1.1	Le dispositif opérationnel prototype	232
9.1.2	L'archive de prévisions numériques	233
9.1.3	Performance opérationnelle des algorithmes	235
9.2	Etude de cas : l'épisode du 8 et 9 septembre 2002 sur les Cévennes	237
9.2.1	Description de l'épisode	237
9.2.2	Prévisions émises le matin du 8 septembre	238
9.2.3	Evolution des prévisions dans les jours précédents	248
9.3	Conclusion	249
	Conclusions	253
	Références bibliographiques	257

ANNEXES

A	L'interface de visualisation des prévisions : RainFAST	267
A.1	Introduction	267
A.2	Mise en route	268
A.3	Affichage des prévisions de précipitations	269
A.3.1	Les cartes de prévisions	269
A.3.2	Détails par groupement	271
A.4	Visualisation des données météorologiques	272
A.4.1	Sélection des données à comparer	273
A.4.2	Options de représentation	275
A.5	Changement de date et de mode de visualisation	276
A.6	Exporter et imprimer les graphiques	276

Avant-propos

**« *La météo est une science qui permet de connaître
le temps qu'il aurait dû faire.* »**

Philippe Bouvard, journaliste

La légende prétend qu'après avoir construit le premier ordinateur en 1946, John von Neumann aurait dit : « *Maintenant, donnez-moi un problème (presque) insoluble (...) !* ». Suite à quoi, il se serait tourné vers la prévision du temps (Moch 2002).

Réels ou imaginaires, ces propos illustrent la complexité des processus atmosphériques et le défi que constitue leur prévision. De plus, parmi l'ensemble des éléments du temps sensible, les précipitations constituent sûrement la variable la plus ardue à prévoir. Il n'est donc pas étonnant que malgré l'existence d'enjeux importants (gestion de la ressource en eau, production d'énergie, inondations, ...), la prévision des pluies reste un domaine encore imparfaitement résolu par les techniques actuelles.

Pour les météorologues, la variable précipitation apparaît plutôt à l'aval du processus de modélisation de l'atmosphère. Pour les hydrologues, chargés de représenter le cycle terrestre de l'eau et de prévoir les débits des rivières, elle est une donnée cruciale en amont de leur domaine de modélisation. Il est donc légitime que tous deux, météorologues et hydrologues, s'attachent à en avoir une bonne prévision et que l'un comme l'autre contribuent à son amélioration.

C'est pourquoi, avec l'espoir d'apporter une aide au prévisionniste aussi bien qu'au modélisateur, nous nous proposons de participer dans ce mémoire à l'étude de l'application du concept d'*analogie* à la prévision des précipitations. Résumé en deux mots, le principe de cette approche est de rechercher, dans une archive météorologique de situations passées, celles qui sont similaires, en terme de circulation synoptique par exemple, à la situation cible de la prévision. En appliquant l'adage « *mêmes causes, mêmes effets* », la connaissance des précipitations observées lors des situations passées reconnues comme analogues permet ensuite d'émettre une estimation probabiliste concernant les précipitations attendues pour cette situation cible.

Cette approche très ancienne, est en fait sous-jacente à toute démarche de prévision météorologique. Son application directe est cependant passée au second plan devant les perspectives qu'ouvrait la modélisation numérique. Les progrès de celle-ci ont d'ailleurs conduit aux modèles météorologiques actuels, qui fournissent des résultats étonnants et constituent un outil indispensable à tout prévisionniste. Néanmoins, l'approche par analogie n'a jamais été complètement

abandonnée et les capacités informatiques actuelles en font une application rapide et légère, susceptible de compléter et de raffiner les résultats de ces modèles. Ceux qui ont continué à en faire usage ont par ailleurs pu constater que, portée par l'amélioration constante des modèles météorologiques, elle ne cessait elle aussi de s'améliorer.

En outre, la récente apparition d'archives météorologiques longues et fiables, les *réanalyses*, permet de s'affranchir du problème de la collecte des données historiques. Ainsi, nous disposons maintenant d'une description des situations météorologiques passées portant sur l'ensemble des variables atmosphériques, ce qui nous autorise à rechercher des analogies plus complètes et nous incite à revisiter l'approche par analogie.

L'utilisation de ces méthodes peut être intéressante dans différents contextes, complémentaires les uns des autres :

- D'une part dans le domaine de la **prévision** : l'approche par analogie a alors vocation à être employée comme une *méthode d'adaptation* des modèles météorologiques. Ceux-ci fournissent des prévisions pour la situation cible, et notamment celles concernant les variables synoptiques, à partir desquelles peut s'effectuer une recherche de situations analogues.

Les modèles météorologiques des services nationaux devant couvrir l'intégralité des processus atmosphériques sur l'ensemble de vastes territoires, il ne peuvent bien sûr pas répondre parfaitement aux besoins spécifiques de chaque utilisateur. L'intérêt réside donc ici dans la possibilité pour ces utilisateurs de raffiner les prévisions générales et de les adapter à leurs objectifs, en produisant par exemple une prévision probabiliste des précipitations sur un bassin versant donné.

- D'autre part dans le domaine de la **simulation** : en s'intéressant à certains types de situations et en recherchant quelles sont les analogies qui « *expliquent* » le mieux les précipitations observées, on peut espérer mettre en évidence certains processus générateurs de précipitations. Cette utilisation diagnostique peut ensuite guider les futurs développements des modèles numériques.
- Mais aussi, pourquoi pas, dans un contexte d'**évaluation** des produits météorologiques. Ces techniques, qui « *court-circuitent* » les modélisations des processus physiques générateurs de précipitations, constituent une sorte de référence de base. A l'aune de cette « *méthode du pauvre* », il est possible de comparer entre eux les résultats générés par différentes paramétrisations physiques des précipitations.

La démarche que nous adoptons ici consiste à rechercher tout d'abord, dans un cadre de *prévision parfaite* (toutes les données sont observées ou analysées), quelles sont les analogies les plus explicatives vis à vis des précipitations, et quelles en sont les caractérisations les plus adaptées. Nous sommes ici dans une démarche de simulation. Ensuite, nous chercherons à évaluer en prévision l'intérêt des algorithmes mis au point. Néanmoins, nous utiliserons indistinctement le terme de prévision dans la suite de ce mémoire, quel que soit le contexte dans lequel nous nous plaçons.

Signalons enfin que notre démarche est probabiliste par nature : nos prévisions s'expriment comme une loi de probabilité d'observer différents cumuls de pluie. L'apport de l'analogie permet de conditionner cette loi le plus possible par la connaissance des spécificités de la situation météorologique cible. Par cet aspect, nous sommes rejoints par les nouvelles pratiques météorologiques qui cherchent à inclure les incertitudes et à les exprimer sous forme de prévisions d'ensemble.

Ce travail est organisé en trois parties :

La **première partie** s'attache à resituer notre approche dans le contexte général de la prévision météorologique et à présenter les outils que nous utiliserons par la suite. Ainsi, le chapitre 1 dresse un panorama succinct des différentes méthodes utilisées en prévision du temps, avant de détailler le principe de l'approche par analogie. Le chapitre 2 présente ensuite les jeux de données dont nous disposons. Nous évoquerons notamment les réanalyses NCEP/NCAR, lesquelles constituent l'« *archive* » météorologique que nous employons. Le chapitre 3 joue un rôle crucial dans ce travail : il s'intéresse à la comparaison objective de prévisions (probabilistes ou déterministes) par rapport aux observations. Il définit ainsi les critères qui nous serviront à reconnaître qu'une analogie est plus pertinente qu'une autre vis à vis des précipitations.

Après la vision très générale de cette première partie, la **deuxième partie** se focalise sur la recherche des algorithmes d'analogie les plus informatifs pour expliquer les précipitations. Dans un premier temps, le chapitre 4 s'intéresse à l'identification de la meilleure variable d'analogie et à l'expression la plus adaptée de cette analogie. Ensuite, le chapitre 5 se préoccupe de la définition du domaine géographique sur lequel il convient de considérer l'analogie. Nous envisageons notamment la possibilité de rendre ce domaine dynamique en le conditionnant en fonction des caractéristiques de la situation cible. Finalement, le chapitre 6 propose de compléter l'analogie en introduisant une caractérisation plus complète des situations cibles. Nous étudions ainsi l'intérêt de l'adjonction d'une analogie portant sur des variables locales, telles que l'humidité, en plus de l'analogie à échelle synoptique basée sur la circulation générale utilisée dans les chapitres 4 et 5.

Enfin, la **troisième partie**, qui se compose de trois courts chapitres, est l'occasion d'études diagnostiques sur les capacités et les limites des algorithmes d'analogies identifiés lors de la deuxième partie. Le chapitre 7 porte ainsi sur la validation de ces algorithmes. Nous y adoptons par ailleurs le point de vue d'un éventuel utilisateur de prévisions pour effectuer une critique des performances obtenues. Nous restons toutefois ici dans un contexte de prévision parfaite pour nos données d'entrée. Le chapitre 8 s'intéresse pour sa part à la sensibilité de ces algorithmes aux caractéristiques de l'archive météorologique. Enfin, le dernier chapitre quitte le domaine de la prévision parfaite pour s'essayer à la prévision opérationnelle, par le biais d'une étude de cas en condition de prévision réelle portant sur l'évènement pluviométrique majeur survenu les 8 et 9 septembre 2002 dans la région des Cévennes.

Première partie

Contexte & Outils

1

L'approche par analogie en prévision météorologique

« Just as relativity eliminated the Newtonian illusion of absolute space and time, and as quantum theory eliminated the Newtonian and Einsteinian dream of a controllable measurement process, chaos eliminates the Laplacian fantasy of long-term deterministic predictability. »

« Tout comme la relativité a détruit l'illusion newtonnienne de l'espace et du temps absolu, et que la physique quantique a détruit les rêves d'Einstein et de Newton de protocoles de mesure sans influence, le chaos détruit la chimère laplacienne de la prédictibilité déterministe à long terme. »
Zeng, Pielke et Eykholt (Zeng et al. 1993)

L'objet de cette thèse est de mettre en relation précipitations et situations météorologiques par l'intermédiaire du concept d'analogie, dans un but de prévisions ou de diagnostics. Ce premier chapitre se propose donc d'introduire et d'expliciter cette notion d'analogie. Mais que le lecteur se rassure, nous allons vite voir que derrière cette appellation obscure se cache tout simplement l'adage bien connu : « *mêmes principes, mêmes effets* » !

Après avoir rappelé brièvement ce qu'un utilisateur potentiel attend d'une prévision de précipitations, nous présenterons les principes des différentes méthodes existantes de prévisions quantitatives de précipitation. Ceci nous amènera tout naturellement à l'approche par analogie, dont nous évoquerons finalement l'historique au sein de la communauté hydrométéorologique grenobloise, et notamment du LTHE.

1.1 Besoins et perspectives d'utilisation

Le choix de la mise en œuvre de l'une ou l'autre des méthodes de prévision de précipitations existantes est naturellement guidé par les besoins de l'utilisateur. Les principales préoccupations de ce dernier sont de trois ordres, et elles dépendent évidemment des enjeux qui président à l'utilisation de prévisions de précipitation :

- l'échéance de prévision, qui dépend de l'échelle de temps de ces enjeux,
- le besoin d'adaptation locale, qui dépend de leur échelle spatiale,
- et la forme revêtue par la prévision, qui dépend de la nature de ces enjeux, mais aussi des capacités de post-traitement de l'utilisateur.

1.1.1 Les échéances de prévision

Le principe d'une prévision météorologique est de prévoir, à partir d'un état mesuré à un instant donné, l'évolution de l'atmosphère pour un instant ultérieur. Une prévision est donc indissociable de l'échéance pour laquelle elle a été émise. Selon un schéma idéal, l'échéance utilisée est définie par les besoins et les contraintes de l'utilisateur ; ce choix impose ensuite le type de prévision à mettre en œuvre. Dans la pratique cependant, les limitations de performance des prévisions entraînent souvent l'adoption d'un compromis entre performances espérées et échéances disponibles.

Le point de vue de l'utilisateur

Hormis l'intérêt que porte tout un chacun aux prévisions météorologiques, notamment en ce qui concerne la planification des activités dominicales, nous pouvons distinguer les principales utilisations qui peuvent être faites des prévisions de précipitations :

- l'organisation de manifestations (épreuves sportives telles que le tournoi de Roland Garros, spectacles en plein air ...), qui nécessite des prévisions de seuil d'intensité de pluie à échelle très locale mais également à échéance très courte,
- la gestion des ressources en eau, qui se contente d'une prévision de lames d'eau moyennées sur de grandes surfaces et de grands pas de temps (24 heures) mais nécessite de longues échéances (de quelques jours à quelques semaines),
- la gestion des crues, qui requiert la prévision de lames d'eau sur des surfaces et des pas de temps dépendants du bassin versant considéré, et ceci sur une large gamme d'échéances (d'une heure à quelques jours).

Le cas de la gestion des crues est l'un des plus complexes. Pour de nombreux bassins versants de taille moyenne (500 à 1000 km²) situés en zone montagneuse, tels les bassins méditerranéens, le temps de réaction à une forte pluie est en général d'à peine quelques heures. Le délai de prévision des crues sur de tels bassins ne peut donc excéder ces quelques heures si l'on ne dispose pas de prévision des précipitations futures. Or, le temps nécessaire au calcul puis à la dissémination de la prévision crée des délais incompressibles, dont il faut tenir compte pour distinguer l'échéance de prévision théorique, du temps dont on dispose en pratique entre la prise de connaissance de la prévision et l'événement prévu proprement dit. Des statistiques établies en 1994 sur les archives

du National Weather Service (USA) ont montré, par exemple, que près de la moitié des alertes de crues émises l'ont été alors que ces crues avaient déjà eu lieu (Zevin 1994).

Idéalement, dans la gestion des crues, les délais souhaités sont de 12 à 24 heures pour une mise en alerte (mise en sécurité des biens et des personnes), de 2 à 3 jours pour une mise en vigilance (mise en astreinte des services d'interventions) voire de plus de 3 jours pour d'éventuelles manœuvres préventives d'ouvrages hydrauliques. Plus les échéances sont longues, moins les échelles de temps et d'espace de la prévision ont besoin d'être fines (Golding 2000). Ainsi, juste avant l'événement, savoir si l'on dispose encore de 2 ou de 3 heures pour évacuer les personnes est crucial, alors que 3 jours avant, une précision temporelle de l'ordre de 12 à 24 heures est suffisante pour préparer les équipes d'interventions.

Le point de vue du prévisionniste

Les phénomènes atmosphériques couvrent une large gamme d'échelles de temps et d'espace, depuis l'effet de la rugosité du sol (quelques dizaines de secondes sur une dizaine de mètres) jusqu'aux ondes stationnaires (plusieurs mois sur des dizaines de milliers de kilomètres), en passant par la frontogénèse (quelques jours sur plusieurs centaines de kilomètres) comme illustré par la figure 1.1 (d'après Orlanski). Le prévisionniste doit donc s'attacher à représenter le mieux possible le phénomène dont l'échelle de temps correspond à l'échéance désirée par l'utilisateur. Comme il n'existe toujours pas de méthode de prévision qui soit capable, de façon opérationnelle, de prévoir l'évolution de l'atmosphère à toutes ces échelles et pour toutes les échéances désirées, on n'active pas les mêmes moyens de prévision selon les échéances que l'on souhaite atteindre.

Il est d'usage courant en météorologie de considérer 4 catégories d'échéances de prévision. Chacune correspond à la prépondérance d'un phénomène atmosphérique particulier et donc à un certain type de méthode de prévision :

- **la prévision immédiate (nowcasting)** : elle répond aux échéances allant de 0 à 3 heures, ce qui correspond au cycle de vie d'une cellule orageuse. Cette technique repose principalement sur les observations issues de l'imagerie radar et sur l'advection des cellules pluvieuses déjà existantes.
- **la prévision à courte échéance** : elle couvre la gamme d'échéances allant de 6 à 72 heures et correspond aux phénomènes rapides de la circulation synoptique, tels que le passage d'un front. C'est actuellement le domaine privilégié des modèles déterministes de prévision numérique (cf. 1.2.2).
- **la prévision à moyenne échéance** : elle se situe dans des échéances allant de 72 heures (3 jours) à 10 voire 15 jours. Elle correspond à la frontogénèse : les fronts qui seront observés ne sont pas encore apparus. Les outils utilisés ici mêlent approche stochastique et modèle numérique déterministe, telle la prévision d'ensemble (cf. 1.2.4.1).
- **la prévision saisonnière** : elle se place au-delà de 15 jours, c'est-à-dire au-delà de la limite théorique de prédictibilité de l'atmosphère. Elle correspond aux phénomènes de très basses fréquences (ENSO, NAO, ...). Le but ici n'est pas de prévoir que tel jour, dans 4 mois, il pleuvra 10 ou 30 mm à tel endroit, ce qui est théoriquement impossible, mais plutôt d'être capable de dire si dans 4 mois on observera des excès ou des déficits de pluviométrie mensuelle par rapport à la normale climatologique. On ne cherche plus à prévoir le temps,

		Echelle climatologique		Echelles planétaire et synoptique		Méso-échelle	Microéchelle	
Longueur	Temps	1 mois		1 jour		1 heure	1 minute	1 s
		Ondes stationnaires	Ondes très longues	Ondes de marée				
10 000 km			Ondes baroclines					
2 000 km				Fronts Cyclones tropicaux				
200 km				Jet de basse couche nocturne Ligne de grain Ondes d'inertie Amas de nuages Ondes de relief				
20 km					Orages Ondes de gravité internes Turbulence en ciel clair Effets urbains			
2 km					Tornades Convection profonde Ondes de gravité			
200 m						Tourbillons secs Thermiques Sillages		
20 m								Panaches Turbulence Frottement au sol (rugosité)

FIG. 1.1: Schéma des échelles de temps et d'espace associées aux différents phénomènes météorologiques (d'après Orlanski 1975, traduit par Juvanon du Vachat 1994).

mais le climat. Dans ce domaine, les meilleurs résultats restent encore ceux obtenus par des approches purement statistiques (1.2.3).

Les échelles temporelles et spatiales des phénomènes atmosphériques sont évidemment liées entre elles. Ainsi, si l'échelle de temps du phénomène atmosphérique considéré est petite, son échelle spatiale l'est aussi. Or, la maille de calcul de la prévision doit y être adaptée : plus cette échelle spatiale est petite, plus la résolution doit être fine. Mais qui dit petite maille de calcul dit aussi quantité plus importante de calculs à effectuer... Il est donc nécessaire dans ce cas de considérer un domaine spatial restreint si l'on veut garder une durée de calcul raisonnable. A l'inverse, si ce phénomène atmosphérique possède une grande échelle de temps, il est nécessaire de considérer un large domaine spatial et d'agréger spatialement les prévisions.

De même, la précision temporelle des prévisions suit la même loi : en prévision immédiate, on attend des résultats au pas de temps horaire ou infra-horaire, cumulés sur 6 ou 12 heures à courte échéance, puis journaliers à moyenne échéance et enfin décennaires voire mensuels en prévision saisonnière.

La question de l'échéance de prévision n'est donc pas une question anodine : elle conditionne non seulement le choix du modèle de prévision à utiliser, mais aussi les résolutions temporelles et spatiales de ces prévisions.

1.1.2 Le besoin d'adaptation locale

La pluie, au même titre que le vent au sol, est une variable qui dépend grandement de l'interaction entre le sol et l'atmosphère, et notamment de la topographie. Or, même si l'atmosphère est un milieu continu présentant une certaine homogénéité, il n'en est rien en ce qui concerne la surface terrestre : celle-ci fait apparaître de fortes discontinuités sur de très courtes distances. Ainsi, il suffit de quelques centaines de mètres pour passer de conditions topographiques de plaine à des conditions de montagne.

Si l'on considère des données de pluie à grande échelle, il est possible de faire apparaître une structure spatiale cohérente. Néanmoins, si l'on considère ces données à une échelle de plus en plus fine, on s'aperçoit qu'au sein de cette structure à grande échelle coexistent de nombreuses irrégularités à petite échelle, dues à des effets locaux. Deux postes pluviométriques séparés d'à peine quelques kilomètres peuvent ainsi présenter d'importantes différences dans leurs mesures, même s'ils sont soumis au même régime de précipitation. Ces effets locaux sont donc importants à prendre en compte pour un utilisateur potentiel qui a besoin de prévisions, soit sur une zone très restreinte (organisation d'événements par exemple), soit sur une extension spatiale plus importante mais dont les contours sont irréguliers (moyenne de précipitation sur un bassin versant).

Toutefois, ces effets locaux sont difficilement simulés par les modèles classiques de prévision numérique. En effet, ceux-ci représentent la surface terrestre par des mailles plus ou moins carrées à l'intérieur desquelles les paramètres de surface sont considérés comme homogènes (figure 1.2, tirée de Karl *et al.* 1989). Actuellement, un modèle global comme ARPEGE (Pailleux *et al.* 2000),

à mailles de taille variable, présente des mailles d'à peu près 20 km de côté à l'endroit où elles sont les plus resserrées. Cette résolution reste insuffisante pour prendre en compte toutes les subtilités de la topographie ou encore pour suivre les contours d'un bassin versant. Les modèles à aire limitée utilisés actuellement en recherche (cf. 1.2.2) peuvent descendre jusqu'à des mailles de 2,5 km de côté, mais cela fait apparaître de nouveaux problèmes dans la représentation des processus physiques, si bien que la qualité des prévisions de précipitations n'est pas systématiquement améliorée par rapport à des modèles possédant des mailles de 10 kilomètres de côté (Ducrocq *et al.* 2002).

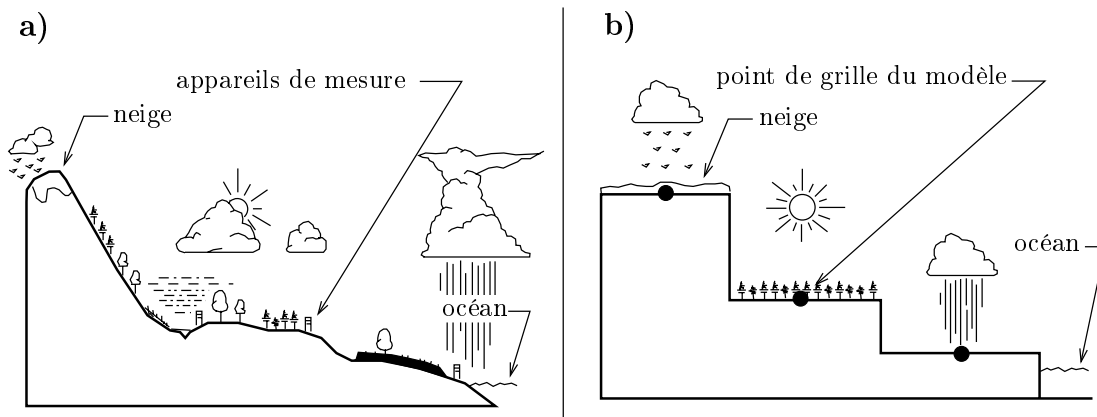


FIG. 1.2: Illustration schématique des différences entre le monde réel (a) et sa représentation (b) par les modèles de prévision numérique du temps (d'après Karl *et al.* 1989).

Parallèlement à cette course à la résolution des modèles numériques, on peut envisager une autre façon de tenir compte des particularismes locaux. En effet, on peut considérer que ceux-ci sont déjà inclus dans les données mesurées à l'endroit considéré. Si, par un effet orographique, cet endroit connaît de plus fortes précipitations que la moyenne alentour, cela apparaît forcément dans les données ; il peut alors être intéressant d'utiliser cette information. Ces considérations, qui ont pour but de fournir des prévisions adaptées à l'échelle locale de l'utilisateur, conduisent aux méthodes d'adaptation statistique des modèles météorologiques (cf. 1.2.4).

1.1.3 L'émission d'une prévision

Il n'y a pas une unique façon d'annoncer une prévision ; celle-ci peut se présenter de différentes manières en fonction de la source dont elle est issue mais aussi et surtout en fonction de l'utilisation qui va en être faite. Ainsi, si l'on prend l'exemple d'une prévision de pluie journalière, celle-ci peut apparaître entre autres :

- sous la forme d'une prévision d'occurrence de pluie pour le grand public : « *demain il va pleuvoir ou il ne va pas pleuvoir* »,
- sous la forme d'une valeur en sortie d'un modèle météorologique numérique : « *demain on observera 27 mm de précipitation à tel endroit* »,
- ou encore sous la forme d'une probabilité au dépassement d'un seuil d'alerte dans un service d'annonce de crue : « *il y a 65% de chance pour que l'on recueille demain plus de 50 mm de pluie sur ce bassin* ».

Prévisions déterministes vs. prévisions probabilistes

On distingue principalement deux formes de prévisions : d'une part les prévisions déterministes, et d'autre part les prévisions probabilistes.

Une prévision déterministe est une prévision qui n'affiche aucune incertitude. Elle fournit, pour une variable à une échéance donnée, une valeur unique à l'utilisateur. Ce type de prévision considère qu'à partir d'un état initial donné de l'atmosphère, il est possible (même si l'on n'en est pas encore capable), de définir de façon déterministe ce que deviendra le système atmosphérique dans le futur. Pour deux états initiaux identiques, l'évolution doit être la même.

Or, les processus atmosphériques sont régis par le *chaos dynamique* : l'évolution temporelle d'un système tel que l'atmosphère est extrêmement sensible à ses conditions initiales (Lorenz 1963, 1993). Etant donné que l'état de l'atmosphère à un instant donné ne peut pas être mesuré de façon exhaustive, il restera toujours une imprécision et une incertitude sur les prévisions. C'est pourquoi on peut considérer, comme l'écrivit Sanders (1963), que la météorologie est de nature stochastique. Les variables atmosphériques gagnent donc à être considérées comme des variables aléatoires dont les lois de probabilité sont conditionnées par ce que nous avons pu mesurer de l'état initial de l'atmosphère.

A ces incertitudes, intrinsèques à la nature de l'atmosphère, s'ajoute le fait que notre connaissance du fonctionnement des processus atmosphériques est imparfaite. Les modèles de prévision ne faisant que résumer l'état de nos connaissances, ils ne représentent en aucun cas l'atmosphère de façon parfaite. Les prévisions déterministes surestiment donc sûrement leurs capacités prédictives et seules les prévisions probabilistes peuvent espérer rendre compte de la réalité, en permettant d'exprimer l'ensemble de ces incertitudes¹

Certes, pour le grand public, les prévisions déterministes sont plus faciles à manipuler. Mais pour un utilisateur final ayant un objectif économique par exemple, les prévisions probabilistes rendent possible l'utilisation des techniques de prise de décision optimale dans un contexte incertain (Massé 1946 ; Katz *et al.* 1982 ; Krzysztofowicz 1998). On peut ainsi montrer que les prévisions probabilistes sont plus utiles, pour l'utilisateur rationnel, que les prévisions déterministes qui ignorent leurs incertitudes, aussi bien dans un contexte théorique (Krzysztofowicz 1983) que dans le cadre d'applications concrètes (Stuart 1982).

Ainsi, dans le cas d'applications économiques impliquant l'utilisation répétée de prévisions météorologique, typiquement dans le cas de la gestion des ressources en eau (barrage, culture, ...), il est licite de chercher à avoir des prévisions utiles « *en espérance* ». Les prévisions probabilistes ont alors tout leur intérêt. L'utilisateur occasionnel de prévision (« *va t-il pleuvoir le jour de mon mariage ?* ») ne bénéficie pas pour sa part de l'effet de moyenne : il préférera la plupart du temps des prévisions déterministes.

¹Différents points de vue peuvent s'affronter ici. Pour notre part, nous ne considérons pas que la réalité est par nature probabiliste, mais plutôt que l'usage des probabilités nous permet de rendre compte des incertitudes liées à notre incapacité à appréhender le monde de façon exhaustive. Le monde n'est pas nécessairement incertain, mais la vision que nous en avons, si.

Le cas de la protection contre les risques naturels est un cas à part. C'est en effet une utilisation systématique des prévisions, mais on ne peut pas se contenter de raisonner en espérance. Un défaut d'alerte n'est en effet pas admissible. Les fausses alertes n'ont, en comparaison, que des conséquences limitées d'un point de vue économique. Par contre, elles influent sur la confiance que porte l'utilisateur aux prévisions et ont un coût caché difficilement quantifiable : à force de trop crier au loup ...

Forme d'une prévision probabiliste

Les prévisions probabilistes, à l'instar des prévisions déterministes, peuvent aussi bien s'appliquer à des variables rangées en un certain nombre de catégories (prévision pluie/non pluie par exemple), qu'à des variables exprimées selon un support continu (prévision de quantité de précipitation par exemple). Dans le premier cas, elles sont exprimées sous la forme de probabilité d'appartenance à chaque catégorie, alors que dans le second cas, elles se présentent généralement sous la forme de *fonctions de répartition*, notées F .

Appliquée à une variable X , $F(x)$ correspond à la probabilité prévue que la variable X dépasse pas la valeur x . Les fonctions de répartition sont donc des fonctions monotones croissantes allant de 0% à 100% (fig. 1.3). D'un point de vue plus mathématique, la fonction de répartition F d'une loi de probabilité est définie par $F(x) = \int_{-\infty}^x f(t).dt$ où f est la densité de probabilité de la loi de probabilité. Si l'on considère la probabilité $P(X < a)$ que la variable aléatoire X tirée de la densité de probabilité f soit inférieure à la valeur a , on a :

$$P(X < a) = \int_{-\infty}^a f(t).dt = F(a) \quad (1.1)$$

Une prévision *déterministe* d'une variable continue peut aussi être représentée sous la forme d'une fonction de répartition. Nous obtenons alors une fonction créneau : ainsi, si la valeur prévue est x^* , alors la fonction de répartition F_{det} est telle que :

$$F_{det}(x) = \begin{cases} 0 & \text{si } x < x^* \\ 1 & \text{si } x \geq x^* \end{cases} \quad (1.2)$$

Dans toute la suite de ce mémoire, nous noterons de telles fonctions de répartition par H_{x^*} .

De même, la valeur observée associée à une prévision peut elle aussi se représenter sous la forme d'une fonction de répartition : si la valeur observée est x^0 , alors la fonction de répartition associée est H_{x^0} (fig. 1.3). Ceci suppose cependant qu'il n'y a aucune incertitude sur l'observation, ce qui est rarement le cas. Normalement, nous devrions en effet considérer que la vraie valeur est distribuée autour de la mesure x^0 . Néanmoins, nous estimons que l'incertitude sur l'observation est significativement plus petite que l'incertitude sur la prévision, ce qui justifie de considérer ici les observations comme des variables déterministes.

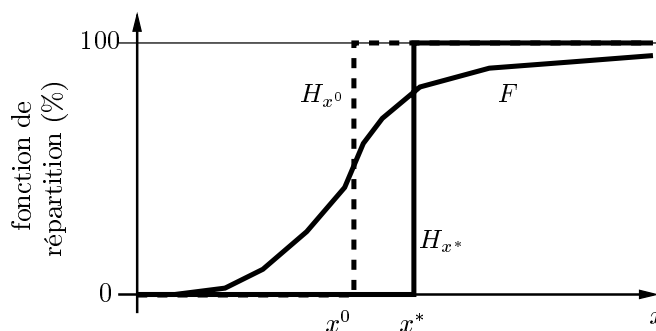


FIG. 1.3: Exemple de fonctions de répartition obtenues pour une variable X par une prévision probabiliste (F), une prévision déterministe (H_{x^*}) et l'observation correspondante (H_{x^0}).

1.2 Les différentes approches

1.2.1 Prévisions objectives et interprétations subjectives

Les différentes méthodes de prévisions présentées ici sont des méthodes objectives. C'est-à-dire que deux opérateurs différents, s'intéressant à la même situation, aboutiront aux mêmes résultats. Ce sont des méthodes automatisées dans lesquelles l'intervention humaine est exclue, par opposition avec les prévisions subjectives, issues de l'interprétation de la situation météorologique par un prévisionniste.

De nombreux auteurs ont été tentés de mettre en compétition prévisions subjectives et prévisions objectives. De telles comparaisons sont cependant dommageables, car elles tendent à donner l'exclusivité à la méthode apparue comme étant la « meilleure », reléguant l'autre au musée des curiosités obsolètes.

Il paraît plus sain de considérer que ces deux approches sont appelées à se compléter et non à s'exclure. C'est pourquoi nous rejoignons l'opinion de Sanders (1963) qui considère une prévision comme un jugement faillible qui peut utiliser toute l'aide issue de processus objectifs dont il peut disposer. Les techniques objectives procurent ainsi un point de référence (probabiliste ou non) que le prévisionniste aiguise ensuite subjectivement grâce à son jugement critique, en utilisant des informations additionnelles. La meilleure méthode apparaît ainsi toujours comme étant la prévision objective interprétée subjectivement. On peut donc dire que, dans son utilisation opérationnelle, la prévision météorologique est un artisanat scientifique.

Même si par la suite nous nous intéresserons de façon exclusive aux aspects objectifs de la prévision, il ne faudra donc pas perdre de vue que toutes les méthodes présentées ici ont vocation à être interprétées par un prévisionniste. Ceci est d'autant plus vrai pour les méthodes probabilistes telles que la méthode des Analogues (*cf.* 1.3).

1.2.2 La modélisation numérique (NWP)

La modélisation numérique déterministe (encore appelée *mécaniste*) consiste à résoudre les équations décrivant le comportement de l'atmosphère. Celles-ci forment un système d'équations non-linéaires qui ne peuvent pas être résolues de façon analytique dans le cas général. La résolution se fait donc de façon numérique et c'est pourquoi le développement de ces méthodes est étroitement lié aux progrès des ordinateurs.

On distingue deux étapes dans la construction d'un modèle numérique de prévision du temps (Coiffier 2000) :

- tout d'abord, il faut établir un système d'équations décrivant l'atmosphère,
- et ensuite il faut numériser ces équations, c'est-à-dire les remplacer par des équations portant sur des variables discrètes, afin de rendre la résolution numérique possible.

Le fait de numériser les équations suppose que l'on transforme un monde continu en une représentation discrète. Plus cette discrétisation est fine, plus le nombre de calculs à effectuer est important, et plus la résolution numérique des équations prend du temps. Si l'on veut garder des temps de calculs compatibles avec une utilisation en prévision, on est obligé de limiter la résolution spatiale de cette discrétisation. Or, comme nous l'avons vu précédemment (cf. 1.1.1), tous les phénomènes météorologiques n'ont pas la même échelle spatiale. Un grand nombre de ces phénomènes se déroule sur des échelles spatiales inférieures à la résolution de la discrétisation. Leur calcul ne peut donc pas se faire via l'intégration numérique des équations : ce sont des processus de sous-maille. Pour tenir compte de la répercussion de leurs effets à plus grande échelle, ces phénomènes sont dits « *paramétrés* » dans le modèle, c'est-à-dire que leur représentation ne se fait pas grâce à la résolution des équations des processus physiques mais par l'intermédiaire de termes correctifs traduisant leurs effets moyens sur les variables du modèle. Ces termes correctifs sont le plus souvent issus d'équations empiriques.

En résumé, on peut donc dire qu'un modèle numérique de prévision du temps est constitué de deux parties :

- Une partie de résolution numérique des équations des processus atmosphériques dont l'échelle spatiale est supérieure à la résolution du modèle. Dans l'état actuel des modèles NWP, ces processus sont en fait liés au mouvement des masses d'air. C'est la partie *dynamique* du modèle.
- Une partie de paramétrisation des effets sur les variables du modèle des processus sous-maille. Cette partie représente surtout des processus micro-physiques et thermodynamiques. C'est la partie *physique* du modèle.

Selon la résolution de la numérisation, la partie physique est donc plus ou moins complète et les paramétrisations plus ou moins importantes. On distingue en général trois types de modélisation, en fonction de la résolution spatiale adoptée :

- **Les modèles de circulation générale (GCM) :** ces modèles sont destinés à faire de la prévision à très longue échéance (plusieurs années) sur l'ensemble du globe. Les échéances atteintes sont tellement importantes que l'on parle alors plutôt de « *simulation climatique* ». Le temps d'intégration étant très long, la résolution spatiale est forcément faible si l'on veut conserver un temps de calcul raisonnable. Actuellement, ce type de modèles fonctionne avec des résolutions horizontales de l'ordre de $1^\circ \times 1^\circ$. Ils ne sont donc aptes à rendre compte que des phénomènes de basse fréquence.
- **Les modèles de prévisions globaux :** ces modèles sont destinés à effectuer des prévisions à court et moyen terme sur l'ensemble du globe. Ils ont des mailles de l'ordre de 50 à 20 km de côté. Ces échelles autorisent l'utilisation d'équations simplifiées grâce à l'hypothèse hydrostatique et imposent la paramétrisation de la convection par exemple.
- **Les modèles à aire limitée (LAM) :** ces modèles sont utilisés pour faire des prévisions à courte échéance sur une région restreinte. Ils utilisent des mailles de 10 à 2.5 km de côté, ce qui les oblige généralement à abandonner l'hypothèse hydrostatique et à représenter de façon explicite la convection. Le domaine considéré par les LAM ne couvrant pas l'ensemble du globe, leurs conditions aux limites doivent être alimentées par des prévisions issues de modèles globaux (on parle de *forçage* du modèle).

En général, les variables liées à la circulation synoptique (partie dynamique du modèle) donnent des prévisions satisfaisantes jusqu'à 3 jours ou plus d'échéance (Atger 2000). Par contre, en ce qui concerne les variables telles que les précipitations, les performances des prévisions quantitatives chutent fortement pour les échéances supérieures à 24 heures.

Il convient de noter que même si certains de ces modèles ne sont pas utilisables de façon opérationnelle pour la prévision, parce qu'ils sont trop gourmands en temps de calculs, ils revêtent une importante capitale au niveau scientifique. Ils permettent en effet de comprendre ou de tester des hypothèses pour déterminer quels sont les processus physiques à l'œuvre dans l'atmosphère.

1.2.3 Les approches statistiques

Les approches purement statistiques cherchent à identifier, à partir d'une archive de données, les relations statistiques qui peuvent exister entre certaines variables (les prédicteurs), observées à un instant t , et la variable cible (le prédicand), observée à un instant $t + \Delta t$. Le but ici n'est pas d'explicitier les processus physiques qui se cachent derrière ces liaisons statistiques mais simplement d'exprimer le mieux possible ces liaisons. Ces approches ont donc une philosophie qui peut sembler totalement opposée à celle des modèles numériques.

Une fois la liaison établie, on s'en sert pour prévoir la variable cible à échéance de Δt à partir de l'observation des prédicteurs au temps t . Pour chaque échéance considérée, il est ainsi nécessaire de formuler une nouvelle liaison statistique.

Mathématiquement, on peut formaliser cette approche de la manière suivante :

$$\begin{cases} \text{calage :} & f_{\Delta t}(x_t^0) = y_{t+\Delta t}^0 + \epsilon \\ \text{prévision :} & y_{t+\Delta t}^* = f_{\Delta t}(x_t^0) \end{cases} \quad (1.3)$$

où :

- $f_{\Delta t}$ est la relation statistique établie pour l'échéance Δt ,
- x_t^0 est le prédicteur observé à l'instant t ,
- $y_{t+\Delta t}^0$ est le prédicteur observé à l'instant $t + \Delta t$,
- $y_{t+\Delta t}^*$ est la valeur prévue du prédicteur pour l'instant $t + \Delta t$,
- ϵ est l'erreur de la relation statistique (que l'on cherche à minimiser).

Ces méthodes présentent l'avantage d'être beaucoup moins coûteuses à développer et à faire fonctionner que les modèles numériques. Jusque dans les années 70, l'extrapolation des champs de géopotentiels observés (par un modèle de Markov) était aussi performante, à 2 ou 3 jours d'échéance, que l'utilisation d'un modèle NWP de l'époque. Les approches purement statistiques ont cependant maintenant tendance à être abandonnées pour les prévisions à court et moyen terme où elles ont été largement supplantées par les modèles numériques. Néanmoins, elles restent encore tout à fait d'actualité pour les prévisions saisonnières. Ainsi, une comparaison menée par Anderson *et al.* (1999) sur la prévision saisonnière en domaine extra-tropical, a montré qu'un algorithme de Canonical Correlation Analysis (CCA), développé par Barnston et Smith (1996) et utilisé comme outil de prévision opérationnel, fournissait de meilleures simulations qu'un modèle numérique, même après traitement statistique des biais de ce dernier. Les méthodes statistiques sont donc encore et toujours utilisées en prévisions saisonnières. Elles n'ont même sûrement pas encore atteint leurs limites, la recherche continuant dans ce domaine et proposant de nouveaux algorithmes (voir par exemple Sharma 2000).

1.2.4 Les approches mixtes : prévision d'ensemble et techniques d'adaptation

Bien que paraissant antinomiques, les approches numériques et statistiques s'avèrent finalement complémentaires et gagnent généralement à être fusionnées. La prévision numérique apporte sa forte capacité prédictive des phénomènes dynamiques alors que les interprétations statistiques permettent de rendre compte à la fois des effets locaux, grâce à l'utilisation de l'information contenue dans les séries de données, mais aussi de l'incertitude liée au chaos dynamique. Nous citons ici quelques exemples de systèmes de prévision mêlant les deux approches.

1.2.4.1 La prévision d'ensemble

La prévision d'ensemble est à considérer comme une utilisation *stochastique* plutôt que *statistique* des modèles numériques. En effet, elle ne fait pas intervenir de liaisons statistiques entre variables ni d'apprentissage sur des données passées.

Du fait du chaos dynamique, l'atmosphère est fortement non linéaire, si bien que les modèles numériques le sont aussi. Une erreur, aussi minime soit-elle, sur les conditions initiales du modèle, peut donc entraîner de très fortes erreurs sur les résultats de prévision. L'idée de la prévision d'ensemble est donc de ne pas se contenter de la prévision reposant sur la meilleure estimation des conditions initiales, c'est à dire le résultat *unique* de la phase d'analyse de toutes les données disponibles, appelé « *l'analyse* ». L'idée est plutôt de réaliser tout un ensemble de prévisions reposant sur tout un ensemble de conditions initiales proches les unes des autres, mais néanmoins différentes (comme deux observations peuvent par exemple l'être si l'on modifie la répartition spatiale des stations de mesures). Chacune de ces prévisions fournit une évolution temporelle appelée une *trace* : elles représentent des évolutions possibles (supposées équiprobables) de l'atmosphère. On obtient ainsi une distribution de prévisions qui permet d'estimer, pour différentes échéances, la probabilité d'occurrence de certains phénomènes météorologiques, selon que ceux-ci se retrouvent dans un nombre plus ou moins important de traces. De cette façon, il est possible d'obtenir soit une prévision probabiliste, soit d'associer un intervalle de confiance à la prévision déterministe issue de l'analyse.

Les différentes conditions initiales utilisées sont obtenues par modifications de l'analyse. Il existe divers protocoles pratiques pour les construire (voir par exemple : Buizza *et al.* 1993 ; Toth et Kalnay 1993), mais tous cherchent à identifier quels sont les états initiaux qui, tout en restant compatibles avec les erreurs de l'analyse, produisent les solutions les plus éloignées les unes des autres (Coiffier 2000).

En pratique, les systèmes opérationnels de prévision d'ensemble, tel celui du Centre Européen de Prévision Météorologique à Moyen Terme (CEPMMT), utilisent une cinquantaine de traces. Cependant, toujours à cause des temps de calcul, le modèle utilisé aujourd'hui pour calculer ces traces est une version simplifiée par rapport au modèle global opérationnel, utilisant une résolution spatiale moindre.

La prévision d'ensemble permet de faire apparaître l'incertitude liée au chaos dynamique de l'atmosphère, en présentant la dispersion possible des solutions envisageables associée à l'incertitude des conditions initiales. Elle ne résout cependant pas les problèmes d'adaptation aux échelles locales. Cette approche peut donc elle-même être conjuguée avec les techniques d'adaptation statistique présentées ci-après.

1.2.4.2 L'adaptation statistique

Les modèles numériques ont, comme nous l'avons vu, des difficultés à représenter les effets locaux. Ils peuvent en outre être entachés de biais, c'est-à-dire d'erreurs systématiques. Afin de les aider à recoller à la réalité, il peut être intéressant de faire appel aux informations contenues dans les séries de données réellement mesurées, celles-ci possédant la connaissance des spécificités locales. On procède ainsi à une *adaptation statistique* des sorties des modèles numériques.

L'adaptation statistique est une variante des approches statistiques *classiques* discutées dans la partie 1.2.3. Comme dans l'approche classique, on procède dans un premier temps au calage

d'une relation statistique entre prédicteurs et prédicands. Puis, cette relation est appliquée pour obtenir une prévision des prédicands à partir des prédicteurs.

La différence est qu'ici, les prédicteurs servant à la prévision, ne sont pas des variables réellement observées, mais sont elles-mêmes prévues par un modèle numérique. Ainsi, pour prévoir la valeur $y_{t+\Delta t}^*$ d'un prédicand, on applique une relation statistique sur les valeurs $\hat{x}_{t+\Delta t}$ des prédicteurs, prévues par un modèle numérique à partir de l'analyse de l'instant t . La projection dans le temps (le passage à une échéance donnée), se fait donc par l'intermédiaire du modèle numérique. On recherche ainsi, dans un contexte d'adaptation statistique, une relation entre prédicteurs et prédicands considérés au même instant.

En fonction de la façon dont on effectue le calage de la relation statistique, on différencie deux contextes d'adaptation (Wilks 1995), comme illustré par la figure 1.4 : la *prévision parfaite* et les *statistiques de sorties de modèle*.

La prévision parfaite (perfect prognosis)

En prévision parfaite, le calage de la relation statistique entre prédicteurs et prédicands s'effectue à partir des données réellement observées. Par contre, cette relation sera ensuite appliquée sur des prédicteurs non pas observés, mais prévus par le modèle numérique :

$$\begin{cases} \text{calage :} & f(x_t^0) = y_t^0 + \epsilon \\ \text{prévision :} & y_{t+\Delta t}^* = f(\hat{x}_{t+\Delta t}) \end{cases} \quad (1.4)$$

où :

- f est la relation statistique établie entre prédicteurs et prédicand,
- x_t^0 est le prédicteur observé à l'instant t ,
- y_t^0 est le prédicand observé à l'instant t ,
- $y_{t+\Delta t}^*$ est la valeur prévue du prédicand pour l'instant $t + \Delta t$,
- $\hat{x}_{t+\Delta t}$ est la valeur prévue du prédicand pour l'instant $t + \Delta t$ par le modèle numérique,
- ϵ est l'erreur de la relation statistique (que l'on cherche à minimiser).

La relation statistique est dans ce cas indépendante à la fois de l'échéance de prévision considérée et du modèle numérique utilisé. Ce type d'approche peut permettre de répondre au besoin d'adaptation locale. Toutefois, si les prévisions des prédicteurs par le modèle numérique présentent un biais, les résultats vont s'en retrouver faussés, car celui-ci n'est pas présent dans les données observées servant au calage de la relation. L'utilisation de cette approche suppose donc que la prévision numérique des prédicteurs n'est pas entachée d'erreurs : qu'elle est parfaite, d'où son nom.

Les statistiques de sortie de modèle (Model Output Statistics, MOS)

Cette approche, introduite par Glahn et Lowry (1972), consiste à effectuer le calage de la relation statistique à partir des prédicteurs prévus par le modèle numérique et non plus à partir de leur observation. Ainsi, si les prévisions présentent un biais, il sera pris en compte lors de

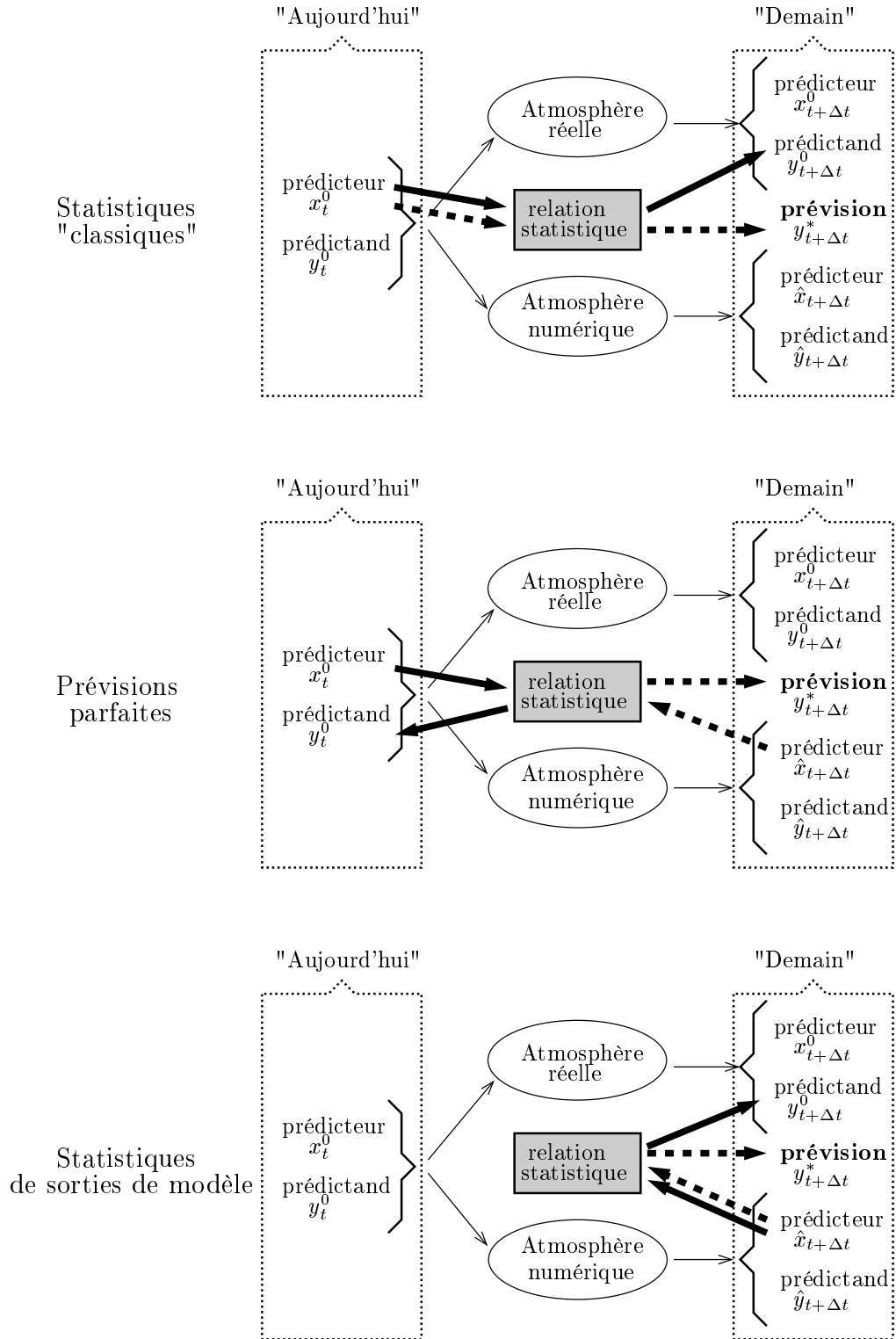


FIG. 1.4: Schéma comparant la technique de prévision statistique "classique" et les deux approches d'adaptation statistique des modèles numériques. Les flèches épaisses pleines indiquent les variables qui interviennent dans le calage des relations statistiques, et les flèches épaisses tiretées indiquent les variables servant à l'application en prévision de ces relations (adapté de Wilks 1995).

l'apprentissage. Les prédictands sont eux évidemment toujours les valeurs observées. Les biais des modèles n'étant pas forcément les mêmes pour toutes les échéances, il convient d'établir une relation différente pour chacune des échéances considérées :

$$\begin{cases} \text{calage :} & f_{\Delta t}(\hat{x}_{t+\Delta t}) & = & y_{t+\Delta t}^0 + \epsilon \\ \text{prévision :} & y_{t+\Delta t}^* & = & f_{\Delta t}(\hat{x}_{t+\Delta t}) \end{cases} \quad (1.5)$$

où :

- $f_{\Delta t}$ est la relation statistique établie entre prédicteurs et prédictand pour l'échéance Δt ,
- $\hat{x}_{t+\Delta t}$ est le prédicteur prévu par le modèle numérique pour l'échéance Δt ,
- $y_{t+\Delta t}^0$ est le prédictand observé à l'instant $t + \Delta t$,
- $y_{t+\Delta t}^*$ est la valeur prévue du prédictand pour l'instant $t + \Delta t$,
- ϵ est l'erreur de la relation statistique (que l'on cherche à minimiser).

Avec cette approche, on peut utiliser la même variable à la fois comme prédicteur et comme prédictand : il est ainsi fréquent d'établir des relations entre précipitations prévues par un modèle et précipitations réellement observées. De cette manière, on corrige le biais du modèle et on peut faire apparaître des effets locaux.

L'approche MOS apparaît plus intéressante que l'approche en prévision parfaite car elle prend directement en compte les biais du modèle. Cependant, cette méthode est plus délicate à mettre en œuvre car elle impose certaines limitations pratiques :

- ① pour pouvoir procéder au calage de la relation statistique, il faut disposer d'une archive suffisamment longue de prévision numérique des prédicteurs,
- ② les relations établies ne sont valables que pour ce seul modèle numérique. En cas de changement du modèle opérationnel, il faut recalculer de nouvelles relations statistiques entre prédicteurs et prédictands.

Le corollaire de ces contraintes est que lorsqu'il y a un changement dans le modèle numérique utilisé, il n'est plus possible d'utiliser d'adaptation MOS tant que l'on ne dispose pas d'une série de prévisions issue du nouveau modèle suffisamment importante. On pourrait s'affranchir de ce délai d'attente en jouant *a posteriori* quelques années passées à chaque changement de version du modèle, mais cela reste prohibitif en temps de calcul. Or, les modèles numériques opérationnels évoluent à un rythme soutenu : ainsi, entre 1992 et 1999, il y a eu en moyenne un changement majeur tous les 6 mois du modèle ARPEGE de Météo-France (Pailleux *et al.* 2000).

Le plus souvent, l'approche MOS est appliquée avec des procédures de régressions linéaires portant sur des prédicteurs représentant la même variable que le prédictand. De tels systèmes sont actuellement utilisés avec succès de manière opérationnelle, notamment dans le cadre de la prévision probabiliste de précipitations (voir par exemple Antolik 2000).

1.3 La prévision par analogie synoptique

La méthode que l'on va présenter maintenant, et qui constitue l'objet de cette thèse, se base sur la recherche d'analogie entre situations météorologiques. Dans son utilisation actuelle, on va voir qu'elle doit être considérée comme une méthode d'adaptation.

1.3.1 Qu'est-ce qu'une *analogie* ?

1.3.1.1 Hypothèse sous-jacente

Le concept d'*analogie synoptique* est très ancien et l'on peut presque dire qu'il remonte aux premiers essais de prévision du temps.

Toute prévision météorologique est basée sur l'hypothèse que l'évolution atmosphérique n'est pas purement aléatoire et qu'elle est régie par un ensemble de lois physiques immuables. Donc, à partir d'un état initial, l'atmosphère ne peut connaître théoriquement qu'une seule évolution possible. Partant de cette hypothèse, plusieurs stratégies de prévision sont apparues :

- On peut chercher à expliciter les lois physiques régissant l'atmosphère, ce qui conduit au développement des modèles numériques (cf. précédemment 1.2.2).
- On peut se contenter d'établir des relations statistiques empiriques entre les états de l'atmosphère pris à des instants différents. On n'explicite alors plus les lois physiques mais uniquement leurs effets : c'est l'approche statistique (cf. 1.2.3).
- On peut enfin considérer que l'atmosphère est un modèle parfait d'elle-même. Donc, si l'on a déjà observé dans le passé une situation similaire (*i.e.* analogue) à la situation qui nous intéresse, leurs évolutions doivent elles aussi être similaires. Ceci constitue l'approche par analogie, ce qui n'est rien d'autre que la mise en pratique de l'adage populaire « *mêmes causes, mêmes effets* », ou plutôt de son corollaire : « *causes proches, effets proches !* ».

Dans la pratique, l'approche par analogie correspond à la démarche naturelle du prévisionniste, qui à la vue d'une situation météorologique se rappelle des situations passées similaires et construit sa prévision en s'aidant des souvenirs qu'il garde de ces situations.

L'analogie est donc un concept très général de reconnaissance des formes : qu'est-ce qui permet de dire que deux individus d'une population ont des "formes" similaires, au sens des variables retenues pour les caractériser ?

Ceci n'est évidemment pas utilisé qu'en météorologie. Ainsi, par le passé, le LTHE a abordé la prévision par analogie via la prévision des avalanches (Bois et Obled 1976 ; Foehn *et al.* 1977 ; Obled et Good 1980). A partir d'une situation nivo-météorologique donnée, dont on veut savoir si elle est propice au déclenchement d'avalanche, on recherche dans une archive des situations passées analogues et on considère si celles-ci ont été plus ou moins fournies en avalanches. Tout le problème réside évidemment dans la définition d'une situation analogue : quelles sont les variables à considérer, et comment le faire ?...

1.3.1.2 Comment identifier une analogue ?

Evidemment, l'analogie parfaite n'existe pas. L'atmosphère est un système trop complexe pour cela. Ainsi, pour caractériser un état atmosphérique de façon exhaustive, il faut connaître, sur l'ensemble du globe, à la fois ses variables dynamiques (pression, vent...), thermodynamiques (température...) et hygrométriques (humidité...), ainsi que les conditions du forçage externe (température des océans, activité solaire...). Le nombre de degrés de liberté de l'atmosphère est donc considérable : rien que pour caractériser un unique champ isobare (500 hPa) sur à peu près un tiers du globe (la partie extratropicale de l'hémisphère nord), il faut considérer de 20 à 35 degrés de liberté selon Van Den Dool (1994) ! Ruosteenoja (1988) estime même la longueur d'archive nécessaire en moyenne pour trouver deux observations identiques de ce seul champ (aux erreurs d'observation près) à 10^{29} années !

Toute l'idée est donc de rechercher des situations passées suffisamment analogues, en se contentant d'une analogie sur quelques variables seulement, astucieusement choisies afin que leurs effets sur les prédictands considérés soient similaires. La première étape consiste donc à définir la **variable d'analogie**, *i.e.* le prédicteur sur lequel on va rechercher la similitude. D'ordinaire, les variables d'analogies sont liées à la dynamique à grande échelle de l'atmosphère (champs de hauteur de géopotential¹ le plus souvent), d'où l'appellation « *d'analogies synoptiques* ».

Ensuite, il faut définir l'extension spatiale de la **fenêtre d'analogie**, c'est-à-dire le domaine spatial sur lequel on recherche l'analogie. En effet, même si l'atmosphère est un système global et que les conditions atmosphériques qui règnent aux antipodes influencent à terme tout le globe terrestre, la propagation des perturbations n'est pas instantanée. Ainsi, si l'on considère des pas de temps ou des échéances suffisamment courts, on peut ne rechercher que des situations localement analogues, comme proposé par Van Den Dool en 1989. Ce même auteur a par ailleurs montré en 1994 qu'en réduisant la taille de la fenêtre d'analogie on augmentait considérablement les chances de trouver de bonnes analogues dans nos archives : les tailles d'archives nécessaires deviennent alors beaucoup plus raisonnable, de l'ordre de 10 à 100 ans. La figure 1.5 présente un exemple de deux situations, dont les champs de pression à 500 hPa, tout en étant différents à l'échelle du globe, sont localement analogues sur une fenêtre de taille réduite.

Enfin, nous avons besoin d'un **critère d'analogie**, *i.e.* un score qui permette de comparer les variables d'analogie de deux situations à l'intérieur du domaine d'analogie. Ce score va permettre

¹Le géopotential Z du niveau z correspond à la variation d'énergie potentielle que subit une masse unité quand elle est portée depuis le niveau de la mer jusqu'à l'altitude z :

$$Z = \int_0^z g(\zeta) \cdot d\zeta$$

où g est l'accélération de la pesanteur. Z s'exprime couramment en « *mètre géopotential* » (mgp) : $1 \text{ mgp} = 9.8 \text{ J} \cdot \text{kg}^{-1}$. On a donc $g \cdot dz = 9.8 \times dZ$. Si l'on considère en première approximation que $g = 9.8 \text{ m}^2 \cdot \text{s}^{-2}$, on peut alors assimiler une variation de 1 mgp à un déplacement vertical de 1 m. Un géopotential correspond donc « *en gros* » à une altitude élaborée.

Une particule d'air se caractérisant souvent par sa pression, on parle couramment du géopotential d'une surface isobare donnée. Sous les latitudes tempérées, on peut fournir les ordres de grandeur suivants concernant l'altitude moyenne des surfaces isobares (atmosphère standard) :

$Z_{1000} \approx 0 \text{ m}$, $Z_{700} \approx 3 \text{ km}$, $Z_{500} \approx 5,5 \text{ km}$ et $Z_{200} \approx 12 \text{ km}$ (niveau de la tropopause)

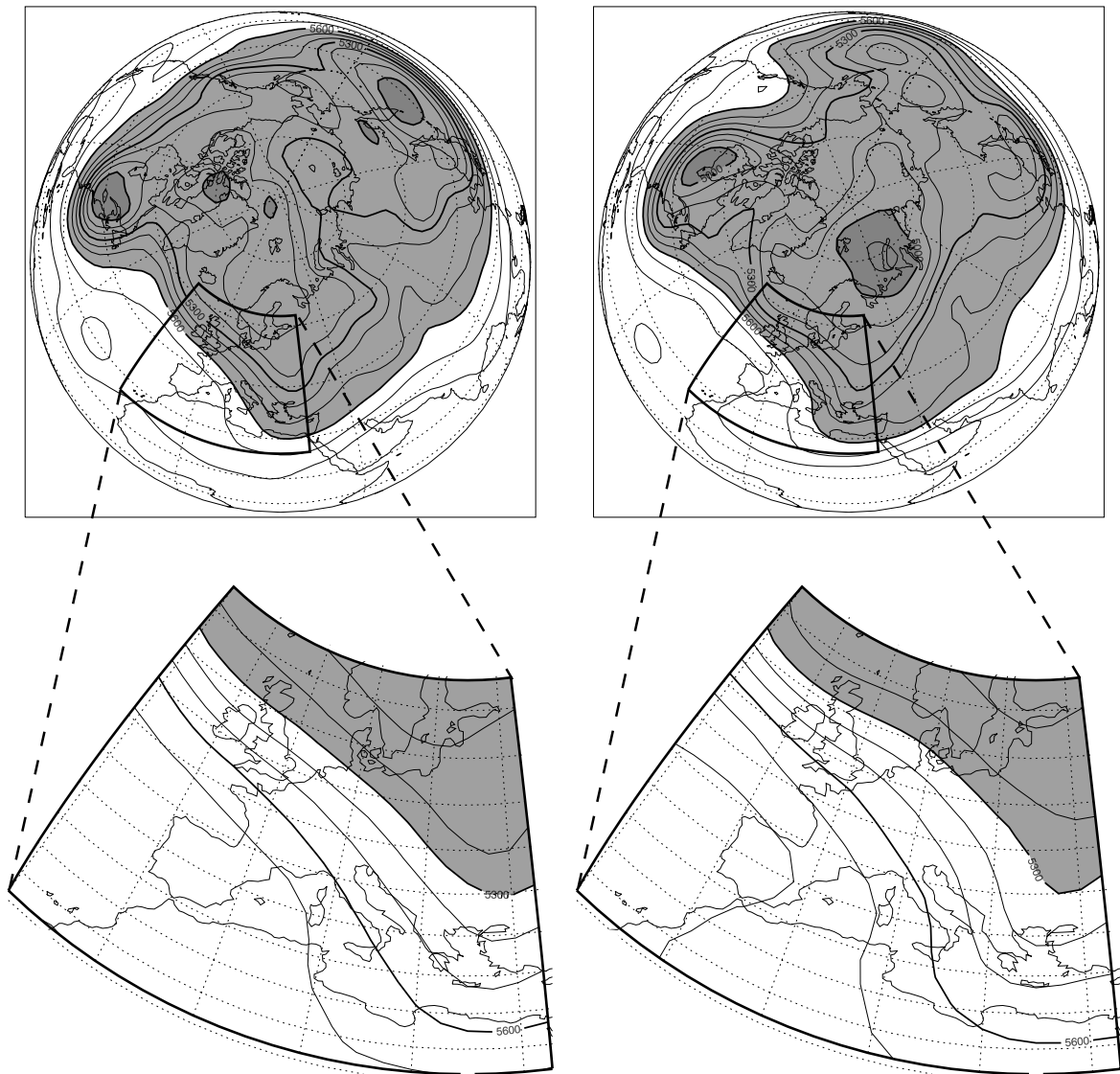


FIG. 1.5: Exemple de deux surfaces isobares à 500 hPa localement analogues : à gauche, situation du 23 janvier 1976 à 00 hTU, et à droite, situation du 3 janvier 1981 à 00 hTU. Les courbes de niveaux sont graduées en mètres.

de classer toute les situations passées, contenues dans une archive météorologique, par rapport à la situation cible : de la plus ressemblante à la plus éloignée.

Une fois toute les journées classées entre elles, on va simplement considérer que la mieux classée est la meilleure analogue à la situation cible (critère des k-plus proches voisins).

1.3.2 Utilisation de l'analogie en prévision

1.3.2.1 Utilisation "classique"

Le concept d'analogie a d'abord été utilisé dans le but de prévoir l'évolution future de champs de pression à partir de leur observation à un instant donné. On considère dans ce cas que deux situations analogues vont *évoluer* de façon similaire. Cette utilisation rentre dans le cadre des prévisions statistiques classiques : le prédicteur est la situation météorologique observée à l'instant t , et le prédicand est la situation météorologique à échéance de $t + \Delta t$.

Malheureusement, l'analogie parfaite n'existe pas et l'on doit se contenter de « *la meilleure analogue de l'archive* ». A cause des archives météorologiques utilisées (d'une durée inférieure à 100 ans), l'écart entre une situation prise au hasard et sa meilleure analogue est très largement supérieur à l'erreur d'une analyse (Lorenz 1969). La situation analogue est donc déjà sensiblement éloignée de la situation observée à l'instant t , et du fait du chaos dynamique, ces deux situations vont vite évoluer de façon divergente. Si l'on considère une analogie locale, on est certes en mesure d'améliorer le degré de ressemblance de la meilleure analogue. Toutefois, cette belle similitude ne peut pas durer bien longtemps, elle non plus : la propagation de l'information située hors de la fenêtre d'analogie va vite entraîner la divergence des deux situations.

L'analogie présente donc de très fortes limitations pour ce qui est de la *projection dans le temps*. A courte et moyenne échéance, cette méthode a très vite été supplantée par les modèles numériques, lesquels se sont rapidement montrés plus efficaces en matière de prévision de variables synoptiques. Ainsi, en 1989, Van Den Dool cite l'exemple de la meilleure paire d'analogues trouvée dans une très longue archive (hauteur isobare à 500 hPa sur la partie extratropicale de l'hémisphère nord, Ruosteenoja 1988), dont l'écart correspond à l'erreur moyenne obtenue par un modèle numérique de l'époque pour une prévision à 4 jours d'échéance.

Néanmoins, en prévision saisonnière, cette utilisation classique des analogues reste relativement performante comparée aux performances des modèles numériques. C'est dans ce domaine que l'utilisation classique des analogues s'est révélée la plus intéressante (Barnett et Preisendorfer 1978 ; Livezey et Barnston 1988 et 1989). Cependant, les récents progrès des modèles GCM semblent indiquer que, là aussi, l'avenir sera plutôt aux méthodes d'adaptations statistiques qu'aux méthodes purement statistiques (Déqué 2003).

1.3.2.2 Utilisation en adaptation statistique

Le concept d'analogie ne peut donc pas être utilisé en prévision en tant que méthode indépendante. La cause est entendue !

Néanmoins, l'analogie n'a pas dit son dernier mot en prévision : elle peut être employée en tant que méthode d'adaptation statistique des modèles numériques. Dans ce contexte, c'est un modèle numérique qui effectue l'extrapolation temporelle, via la prévision des variables synoptiques. La recherche d'analogie est utilisée ici pour relier la situation synoptique prévue à des variables que les modèles numériques ont plus de mal à prévoir, telle la précipitation. Peu importe que

la situation analogue évolue ensuite de la même façon ou pas, puisque l'évolution synoptique n'est pas prévue par analogie. La figure 1.6 schématise les différences entre cette approche et la classique.

L'hypothèse que l'on fait ici est que deux situations synoptiques similaires doivent engendrer des effets locaux similaires. Cette hypothèse avait été émise dès 1956 par Lorenz (cité par Woodcock 1980), qui suggérait de construire les régressions linéaires des approches statistiques sur des échantillons de situations analogues. Son idée est en effet que le fait de considérer une linéarité seulement « *locale* » en terme de situation météorologique doit permettre de conserver et de suivre le caractère globalement non-linéaire de l'atmosphère. En gros, cela revient à faire une approximation linéaire par morceaux des processus atmosphériques.

Il est à noter que l'approche par analogie est notablement différente d'une simple stratification des données en types de temps. En effet, les découpages en types de temps sont effectués une fois pour toutes au moment de la mise en place de la méthode, et les relations statistiques utilisées sont alors fixées. Celles-ci ne changent pas d'une prévision à l'autre du moment que l'on se retrouve avec un même type de temps. Cette approche est donc assez rigide et figée. A l'inverse, l'analogie fonctionne selon un principe de plus proches voisins, qui est beaucoup plus flexible. La relation statistique se base donc pour chaque prévision sur un ensemble différent de situations analogues. Cette relation change ainsi pour chaque prévision, et c'est pourquoi, il arrive que l'on fasse référence à l'adaptation par analogie en parlant d'adaptation *dynamique*, par opposition aux méthodes classiques d'adaptation. En général, cette adaptation dynamique, par analogie, s'avère supérieure à une adaptation statique, par type de temps ou autre, comme l'ont par exemple montré Bois *et al.* (1981) et Obled et Bois (1982) dans le cadre de la prévision d'insolation. D'autres applications de l'analogie en tant que principe d'adaptation statistique des modèles numériques, menées aussi bien sur les précipitations (Duband, 1981) que sur les températures (Woodcock 1980) ont aussi montré l'intérêt de cette approche.

Rappelons ici que les prévisions des variables du temps sensible (températures, précipitations, ...) par les modèles NWP sont bien moins performantes que celles des variables synoptiques, surtout en terme d'échéance de prévision. Les premières sont en effet issues de la partie physique des modèles NWP, alors que les secondes sont issues de leur partie dynamique. Or, l'adaptation synoptique par analogie se base essentiellement sur les prévisions des variables synoptiques. En quelque sorte, on peut ainsi dire qu'elle « *court-circuite* » la partie physique des modèles NWP. On peut donc aussi s'attendre à ce que les performances de l'adaptation par analogie soient plus fortement liées aux performances des prévisions des variables synoptiques, permettant ainsi de fournir des prévisions des variables du temps sensible pour des échéances plus importantes que celles autorisées par la seule utilisation directe des modèles numériques.

Bien sûr, même en prévision parfaite (*i.e.* avec des variables synoptiques observées et non pas prévues), l'adaptation par analogie ne fournit pas une prévision parfaite du prédicand. Il est même possible que, sur les premières échéances, les prévisions numériques de précipitations soient plus performantes que les prévisions obtenues par adaptation. Cependant, la qualité de la prévision des variables synoptiques se dégradant moins vite que celle des variables sensibles, on

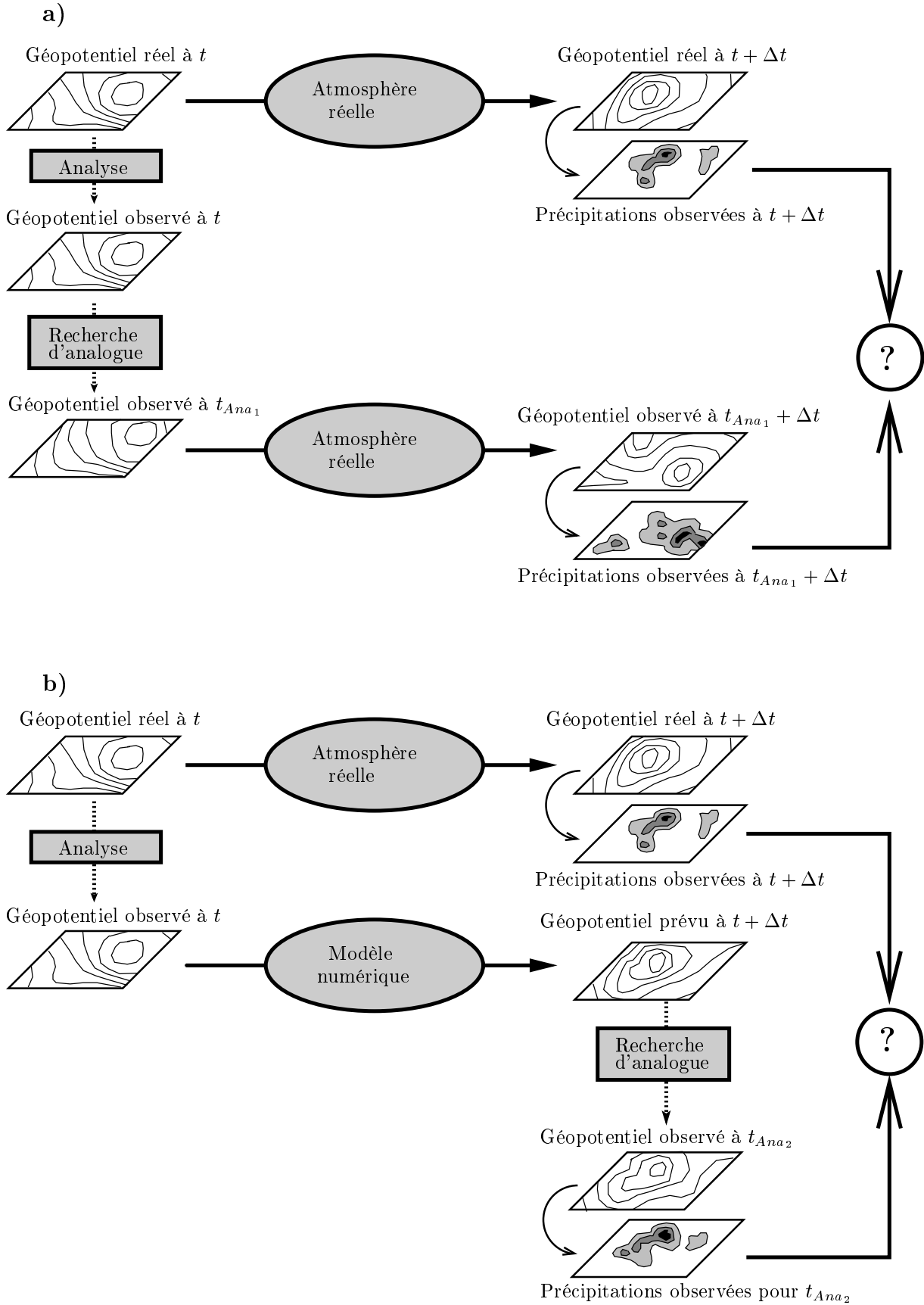


FIG. 1.6: Comparaison de l'utilisation de l'analogie : (a) selon l'approche classique, (b) selon une approche d'adaptation de modèle numérique.

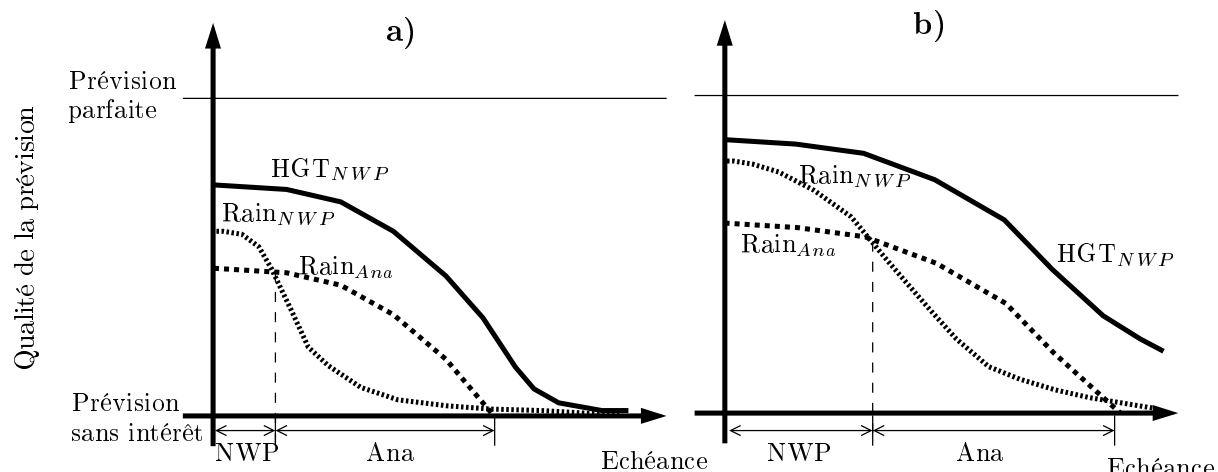


FIG. 1.7: Schématisation des performances attendues des prévisions de géopotentiels (HGT_{NWP} , trait plein) et de précipitations ($Rain_{NWP}$, trait pointillé) d'un modèle numérique comparées à celles attendues sur les prévisions de précipitation par adaptation par analogie ($Rain_{Ana}$, trait tireté). Le panneau (b) présente les performances attendues après amélioration du modèle numérique par rapport à l'état considéré dans le panneau (a).

peut raisonnablement espérer qu'il existe toute une gamme d'échéances pour laquelle l'adaptation par analogie se révélera plus efficace que la prévision directe par modèles NWP (fig. 1.7 a).

Avec l'amélioration constante des modèles numériques, on peut s'attendre à ce que la prévision des variables du temps sensible s'améliore aussi, supplantant l'adaptation par analogie là où cette dernière était encore supérieure. Est-ce à dire que l'adaptation n'a qu'un intérêt temporaire, en attendant les progrès de la modélisation numérique ? Il nous semble que non, au contraire ! En effet, si la qualité de la prévision des variables sensibles augmente, celle des variables dynamiques aussi, et donc par là même, celle de la prévision par adaptation. La gamme d'échéances pour laquelle la prévision par adaptation est plus performante que la prévision directe par modèle NWP sera certes décalée dans le temps, mais elle existera toujours (fig. 1.7 b).

1.3.2.3 Les archives nécessaires

La principale limitation de l'application de l'analogie réside dans la nécessité de disposer d'importantes archives météorologiques. Celles-ci doivent regrouper à la fois les variables d'analogie mais aussi les prédictands qui nous intéressent. Nous reviendrons plus longuement sur ce point dans le chapitre 2 mais il est important de garder à l'esprit leur présence derrière toute recherche d'analogie. De leur qualité (variété des variables, précision des données, profondeur des archives, ...) dépendront les performances des prévisions émises.

1.3.3 Etat de référence (travaux Guilbaud, 1997)

Nos travaux s'inscrivent évidemment dans une certaine continuité. Ainsi, un système opérationnel de prévision de précipitations basé sur une adaptation par analogie du modèle numérique ARPEGE de Météo-France fonctionne actuellement au sein de la Division Technique Générale de

EDF. Il va constituer le point de départ du travail présenté ici et formera ce que l'on appellera par la suite la méthode de référence. Ce système fournit actuellement des prévisions journalières de précipitations sur une cinquantaine de groupements pluviométriques (de 500 à 2000 km²) répartis principalement sur le quart Sud-Est de la France.

1.3.3.1 Historique

A partir des travaux sur les relations entre analogie synoptique et quantité de précipitations recueillies au sol, menés par Duband en 1970, un système de prévision opérationnel fut installé dès le début des années 70 au sein de EDF (Duband 1971, 1974). Ce système était à mi-chemin entre une application purement statistique et une adaptation : la projection dans les temps futurs était alors effectuée par l'application d'équations d'autocorrélation sur les composantes principales des champs de pression. On peut même dire que l'on était alors en présence d'une méthode d'adaptation statistique de modèles statistiques ! Elle était basée sur :

- la recherche d'analogues à partir de l'analyse en composante principale du champ de pression 700 hPa à 00 hTU du jour courant,
- l'utilisation d'un domaine d'analogie couvrant l'Europe de l'Ouest et le proche Atlantique,
- et le calage d'une équation de régression linéaire multiple reliant les précipitations observées lors des journées analogues, aux composantes principales des champs de pression au sol et de température à 700 hPa.

Ce système fournissait des prévisions jusqu'à 48 heures d'échéance, grâce à l'extrapolation statistique des champs de pression.

Dans les années 80, les progrès réalisés sur les modèles numériques ont permis l'utilisation fructueuse de leurs prévisions concernant les variables synoptiques jusqu'à plusieurs jours d'échéance. Le système qui fonctionnait à EDF a alors été modifié afin d'utiliser les prévisions de champs de pressions fournies par un modèle NWP de Météo-France. Il est alors devenu un vrai système d'adaptation de modèle numérique. Le système subit à cette occasion d'autres modifications :

- La recherche d'analogues s'effectue en deux temps :
 - ① on recherche d'abord, à partir des composantes principales du champ de géopotentiel 700 hPa, les journées les plus proches au sens d'une distance euclidienne (critère de proximité).
 - ② et ensuite, au sein des journées sélectionnées par le critère de proximité, on retient les journées les plus ressemblantes grâce à un critère de corrélation portant sur les composantes principales des champs de pression à 700 et 1000 hPa ainsi que du champ de l'épaisseur séparant ces deux géopotentiels (critère de forme).
- L'utilisation des équations de régression entre température et précipitation a été abandonnée au profit de l'émission d'une prévision probabiliste de précipitation. Celle-ci est obtenue par interpolation de la distribution empirique cumulée des précipitations observées lors des journées analogues. La prévision est exprimée par l'intermédiaire de 3 quantiles au non-dépassement : les quantiles 20, 60 et 90 %.

Vers le milieu des années 80, avec l'évolution des performances des modèles numériques, les échéances de prévisions ont été augmentées jusqu'à 72 puis 96 heures.

Enfin, S. Guilbaud a revisité ce système en 1997 et a obtenu plusieurs avancées significatives. Les résultats de ses travaux (Guilbaud 1997 ; Guilbaud et Obled 1998 ; Obled *et al.* 2003) ont permis une mise à jour du système opérationnel à la fin des années 90. C'est ce qui va former notre système de référence.

1.3.3.2 Description de la méthode de référence

Considérons une journée cible C pour laquelle on souhaite émettre une prévision. Cela peut être soit la journée courante (journée J), soit une journée située à une échéance plus lointaine de k jours (journée $J + k$). Le prédictand de la méthode de référence est le cumul journalier de précipitations de 6 hTU le jour C à 6 hTU le jour $C + 1$, sur un groupement pluviométrique donné.

On procède en deux temps :

- ① On recherche tout d'abord un ensemble de situations analogues à la situation prévue pour le jour C par un modèle numérique.
- ② Puis à partir des précipitations observées lors de ces situations analogues, on émet une prévision pour le groupement considéré.

Recherche des situations analogues

En premier lieu, cette méthode prend en compte les effets saisonniers. En effet, les échanges radiatifs (insolation, état du couvert végétal, présence de neige au sol, ...) ne sont pas les mêmes d'une saison à l'autre. Ainsi, deux situations, même si elles présentent des champs de pressions similaires, risquent fort de ne pas produire les mêmes effets locaux si l'une est observée en été et l'autre en hiver, comme cela a été souligné par Lorenz (1969). Pour remédier à ce problème, on ne s'autorise à rechercher des situations analogues que dans une période de l'année qui encadre la date de la journée C de plus ou moins 2 mois (fenêtre temporelle glissante).

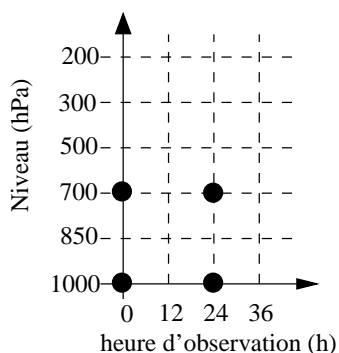


FIG. 1.8: Représentation schématique de la répartition spatio-temporelle des champs de géopotentiels sur lesquels porte l'analogie.

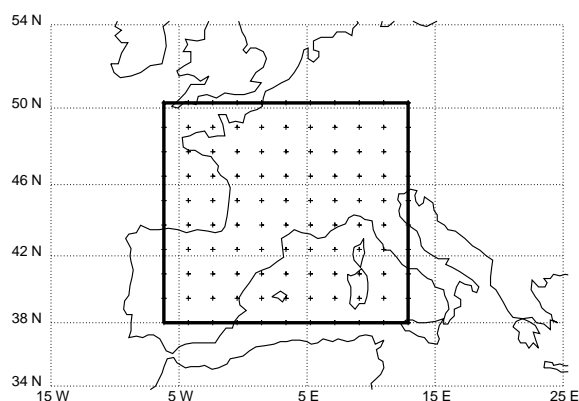


FIG. 1.9: Fenêtre d'analogie utilisée dans la méthode de référence et localisation des points de grille sur lesquels sont représentés les champs géopotentiels.

L'analogie est basée sur 4 champs de géopotentiels issus de prévisions numériques : les champs 700 et 1000 hPa, considérés à 00 et 24 hTU (fig. 1.8). Ce choix permet d'encadrer la période de cumul des précipitations et donne une vision à la fois spatiale et temporelle de l'atmosphère.

L'utilisation des composantes principale a été abandonnée : ces 4 champs de géopotentiels sont maintenant utilisés directement sous la forme de points de grille. Le critère d'analogie employé est le score $S1$ (Teweless et Wobus 1954) que nous expliquons en détail plus tard (cf. 4.1.1.2). En deux mots, ce score est un critère de forme qui, à partir de deux cartes de géopotentiels, compare les champs de vent géostrophique associés. Il n'y a plus maintenant de critère de proximité utilisé. La fenêtre d'analogie considérée s'est pour sa part notablement réduite : elle s'étend actuellement de 6.2°W à 12.9°E et de 38.0°N à 50.3°N (fig. 1.9).

La distance d'analogie totale $DA(J_{ana})$ d'une journée J_{ana} contenue dans l'archive météorologique est au final calculée comme étant la moyenne arithmétique des scores $S1$ par rapport à la journée cible C , calculés sur chacun des 4 champs d'analogie :

$$\begin{aligned}
 DA(J_{ana}) = \frac{1}{4} \cdot & \left[S1 \left(HGT_{(700;00)}^C; HGT_{(700;00)}^{J_{ana}} \right) \right. \\
 & + S1 \left(HGT_{(700;24)}^C; HGT_{(700;24)}^{J_{ana}} \right) \\
 & + S1 \left(HGT_{(1000;00)}^C; HGT_{(1000;00)}^{J_{ana}} \right) \\
 & \left. + S1 \left(HGT_{(1000;24)}^C; HGT_{(1000;24)}^{J_{ana}} \right) \right] \quad (1.6)
 \end{aligned}$$

où $HGT_{(700;00)}^J$ représente le champ de géopotentiel à 700 hPa considéré à 00 hTU de la journée J .

Ainsi, chaque situation de l'archive dont la position dans l'année est située à plus ou moins 2 mois de la journée cible se voit attribuer une distance d'analogie qui reflète son degré de similitude avec la situation cible.

L'émission de la prévision

A partir des scores d'analogies, toutes les situations de l'archive peuvent être classées, de la plus ressemblante à la plus éloignée. Les 50 situations présentant la plus faible distance d'analogie sont alors retenues comme étant les situations analogues à la situation cible. Ce choix de 50 situations résulte d'une optimisation heuristique, à l'instar des choix précédents.

On considère ensuite les cumuls de précipitation observés lors de ces situations analogues sur le groupement pluviométrique d'intérêt. Ces cumuls sont classés du plus petit au plus grand afin de tracer leur distribution empirique. Cette distribution, conditionnée par le fait que l'on a des situations synoptiquement analogues, forme la fonction de répartition prévue : c'est la prévision de précipitation.

Afin de condenser l'information contenue dans cette fonction de répartition, la prévision est émise sous la forme d'un quantile au non-dépassement médian : le quantile 60%, assorti d'un intervalle de confiance : les quantiles 20 et 90% (fig. 1.10).

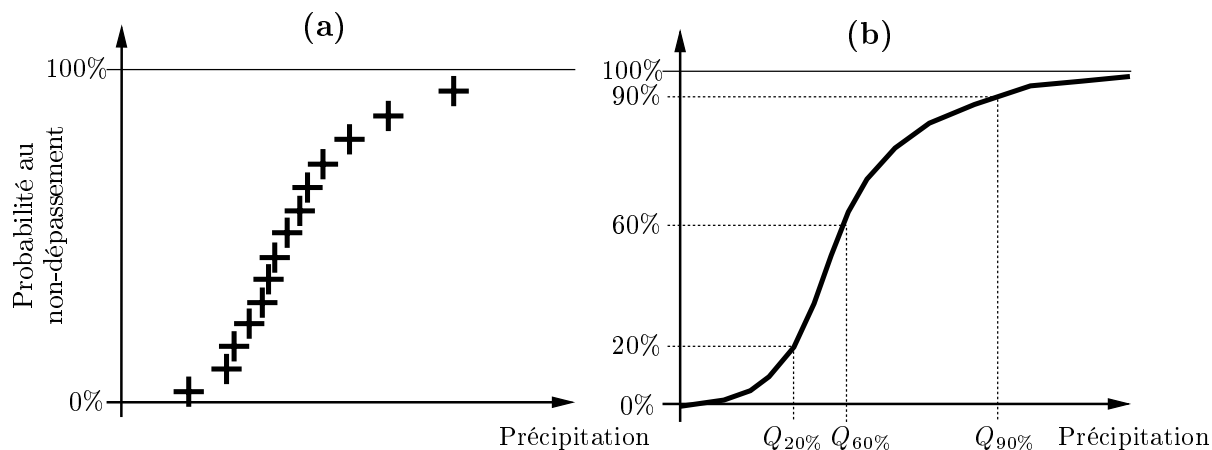


FIG. 1.10: Exemple d'émission d'une prévision sur un groupement pluviométrique : **(a)** distribution empirique des précipitations observées lors des journées analogues ; **(b)** fonction de répartition prévue des précipitations et quantiles 20, 60 et 90% associés (notés $Q_{20\%}$, $Q_{60\%}$ et $Q_{90\%}$).

1.3.3.3 Quelques exemples de prévisions obtenues

Prévision parfaite

La figure 1.11 présente un exemple de prévisions réalisées en prévision parfaite (*i.e.* avec des prédictands observés) sur le bassin du Tanaro. Il apparaît que les prévisions co-fluctuent de manière globalement satisfaisante avec les observations. Elles arrivent par ailleurs à faire une bonne discrimination entre journées pluvieuses et non-pluvieuses. On constate quelques fausses alarmes, pour lesquelles des précipitations sont attendues mais non observées (8 et 10 novembre), mais aucun défaut d'alarme. On peut cependant déplorer que la fourchette d'incertitude produite par la prévision a tendance à devenir très large. Ainsi, pour la journée du 5 novembre, le quantile 20% n'atteint pas 20 mm alors que le quantile 90% dépasse les 100 mm : le prévisionniste est certes informé qu'il y a un risque de pluie importante, mais 20 mm de pluie n'ont pas les mêmes conséquences que 100...

Il est rassurant de voir que les prévisions s'écartent notablement de la climatologie des précipitations sur l'archive. En effet, considérer la distribution des précipitations sur un ensemble de situations analogues revient à effectuer un échantillonnage de la climatologie, conditionné par la ressemblance des situations synoptiques. S'il n'y avait pas de relations entre conditions synoptiques et précipitations observées, on réaliserait un simple échantillonnage aléatoire de la climatologie : nous obtiendrions donc chaque jour des prévisions proches de celle-ci. Or, les prévisions se différencient très nettement de la climatologie, et ceci aussi bien dans les cas de présence que dans les cas d'absence de précipitations. L'échantillonnage n'est donc absolument pas aléatoire, ce qui démontre la pertinence des analogies effectuées.

Utilisation en opérationnel

La figure 1.12 présente un exemple de prévisions opérationnelles obtenues sur le bassin de la Toce (Italie) pour trois échéances différentes. Cette fois-ci, la recherche d'analogues est utilisée comme une méthode d'adaptation d'un modèle numérique : le modèle ARPEGE de Météo-France

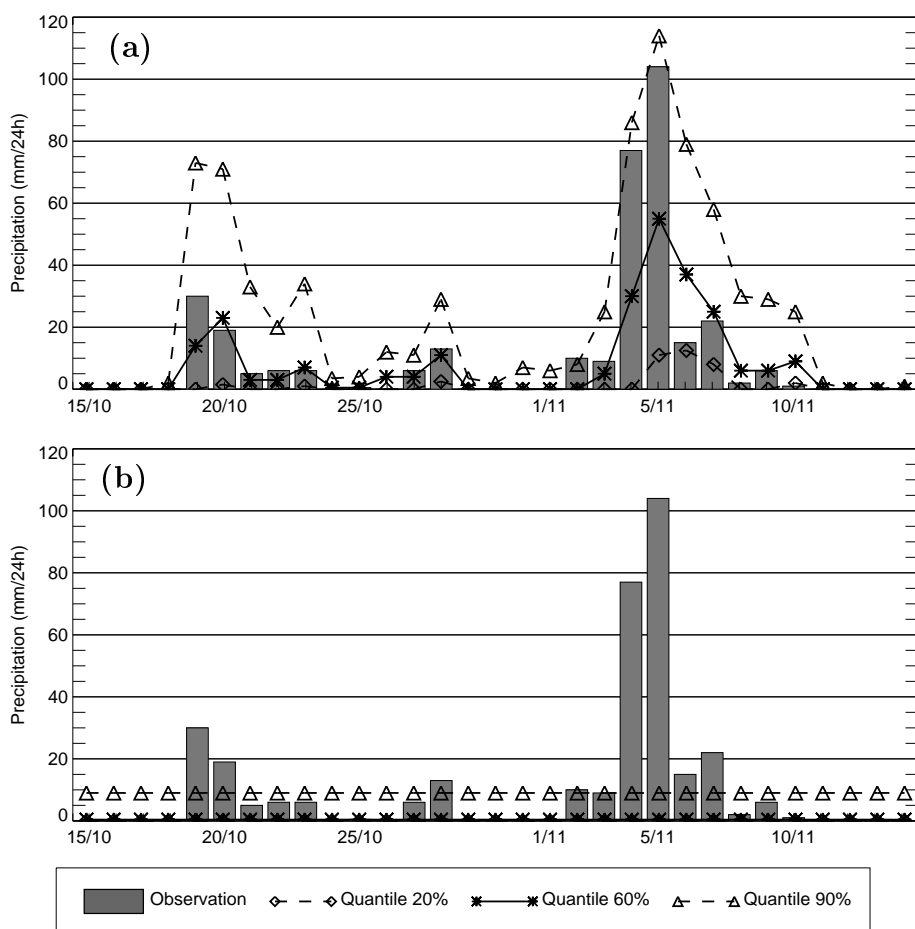


FIG. 1.11: Comparaison entre les précipitations observées et prévues du 15 octobre au 15 novembre 1994 sur le bassin versant du Tanaro (Italie) : (a) prévisions obtenues par la méthode de référence en prévision parfaite, (b) prévisions selon la fréquence climatologique.

fournit les prévisions pour les premières 72 heures, et le modèle du CEPMMT pour les échéances de prévision suivantes. Donc, aux erreurs dues à la recherche d'analogie (erreur de prévision parfaite), s'ajoutent les erreurs dues à la prévision des variables d'analogies par les modèles NWP.

On constate que pour l'échéance J (fig. 1.12 a), les prévisions se comportent globalement comme en prévision parfaite : les observations sont bien cernées par la prévision. On déplore certes quelques fausses alarmes (15 septembre par exemple), mais aucun défaut d'alerte. Ces prévisions s'avèrent quantitativement très intéressantes pour un prévisionniste.

Avec l'augmentation de l'échéance (fig. 1.12 b et c), les prévisions perdent de leur précision, surtout en ce qui concerne les faibles précipitations, pour lesquelles elles deviennent même assez chaotiques à J+5. Cependant, même à J+5, tous les forts épisodes sont effectivement bien prévus, à un décalage d'un pas de temps près (24 heures) : la « signature » synoptique de ces événements est déjà présente dans les prévisions numériques. Les prévisions de précipitations par analogie restent donc qualitativement informatives pour un prévisionniste, même à de telles échéances.

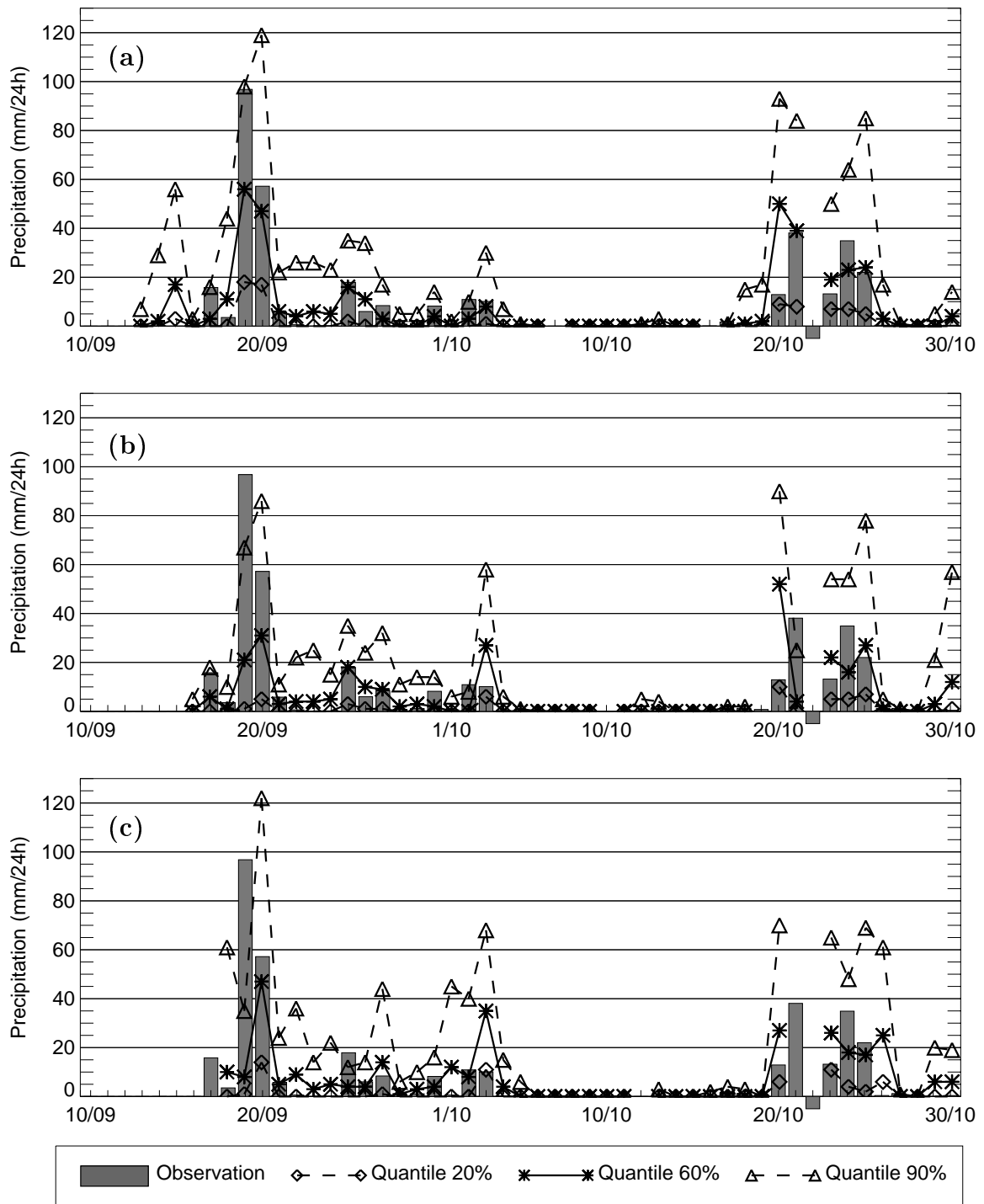


FIG. 1.12: Comparaison entre les précipitations observées et prévues du 10 septembre au 30 octobre 1999 sur le bassin versant de la Toce (Italie) : (a) prévisions émises pour l'échéance J (*i.e.* 0 heure avant le début de la période de cumul), (b) prévisions émises pour l'échéance J+3 (*i.e.* 72 heures avant le début de la période de cumul), (c) prévisions émises pour l'échéance J+5 (*i.e.* 120 heures avant le début de la période de cumul).

Note : les lacunes dans les prévisions correspondent à des problèmes de transmission des prévisions de géopotentiels. De même, les observations négatives (22 octobre) correspondent à des données manquantes.

1.3.4 Pourquoi une prévision probabiliste ?

Arrivé à ce stade, on est en droit de se poser une question importante concernant cette méthode de prévision : pourquoi considérons-nous 50 situations analogues et non pas seulement la meilleure ?

La première réponse est que c'est tout simplement un résultat heuristique. En testant cette adaptation sur 40 années de données et en s'appuyant sur un nombre d'analogues compris entre 1 (la meilleure) et 100, S. Guilbaud a montré que les meilleurs résultats étaient obtenus en utilisant entre 30 et 60 analogues.

On peut cependant essayer de fournir une explication plus théorique à cette question en considérant le principe même de l'adaptation par analogie. En effet, si nous étions capable de comparer deux états de l'atmosphère de façon précise, en prenant en compte exhaustivement toutes les variables d'états, alors la meilleure prévision serait fournie en considérant seulement la meilleure analogue.

Toutefois, pour des raisons pratiques, mais aussi parce que notre connaissance est limitée, nos critères d'analogies ne reposent que sur un petit ensemble de variables synoptiques $\{V_{syn}\}$: nous ignorons l'ensemble des autres variables caractéristiques de l'atmosphère et notamment toutes les variables plus représentatives des échelles locales $\{V_{locales}\}$. La précipitation $P_{obs}^{S^0}$, observée lors d'une situation météorologique S^0 sur un bassin versant donné, dépend donc à la fois des variables $\{V_{syn}\}$ et des variables $\{V_{locales}\}$. En première approximation, nous considérons que ces deux composantes s'ajoutent simplement :

$$P_{obs}^{S^0} = P_{syn}^{S^0} + P_{locales}^{S^0} \quad (1.7)$$

où : $\left| \begin{array}{l} P_{syn}^{S^0} \text{ est la partie de } P_{obs}^{S^0} \text{ due aux variables synoptiques } \{V_{syn}\}, \\ \text{et } P_{locales}^{S^0} \text{ est la partie de } P_{obs}^{S^0} \text{ due aux variables locales } \{V_{locales}\}. \end{array} \right.$

Considérons maintenant une situation S^1 analogue au sens des variables synoptiques $\{V_{syn}\}$ à la situation S^0 . La précipitation $P_{obs}^{S^1}$ observée sur le même bassin versant peut donc elle aussi se décomposer en deux termes :

$$P_{obs}^{S^1} = P_{syn}^{S^1} + P_{locales}^{S^1} \quad (1.8)$$

Du fait de l'analogie entre S^0 et S^1 , nous avons $P_{syn}^{S^1} = P_{syn}^{S^0}$. Par contre, $P_{locales}^{S^1}$ n'est qu'un échantillon parmi d'autres de ce qui peut être induit par la variabilité des conditions locales et nous avons donc généralement $P_{locales}^{S^1} \neq P_{locales}^{S^0}$.

Si nous disposions d'une archive météorologique infinie, nous serions en mesure de trouver une infinité de situations (S^i) rigoureusement analogues à S^0 , en terme de variables synoptiques $\{V_{syn}\}$. Nous aurions alors $P_{syn}^{S^i} = P_{syn}^{S^0}$, si bien que la distribution des ($P_{obs}^{S^i}$) correspondrait à la distribution conditionnelle de la partie des précipitations due à la variabilité des variables locales $\{V_{locales}\}$ pour une situation synoptique fixée. Dans ces conditions, le fait d'émettre une prévision probabiliste se basant sur un nombre important d'analogues permet d'approcher cette

distribution conditionnelle : on exprime ainsi l'incertitude due à la non-prise en compte de l'état des conditions locales dans l'analogie.

En pratique, nos archives ne sont pas infinies et il ne nous est pas possible de trouver d'analogues parfaites, même sur un nombre restreint de variables. C'est pourquoi on se contente de sélectionner un nombre fini de situations « *suffisamment* » similaires. Plus le nombre retenu est important, meilleur est l'échantillonnage des conditions initiales possibles. Cependant, les situations sélectionnées étant alors moins similaires, cet échantillonnage est aussi moins conditionné par la situation synoptique. Il y a donc un compromis à trouver, qui est fonction de la taille de l'archive disponible. Ceci peut être illustré en considérant deux cas extrêmes :

- Si l'on considère seulement la meilleure analogue, alors on est presque sûr que les parties *synoptiques* des précipitations prévues et observées sont similaires. Toutefois, il est aussi presque certain que cette meilleure analogue diffère de S^0 en terme de conditions locales, si bien qu'au final, la prévision sera différente de l'observation. Or, nous ne sommes pas capables, de cette manière, de faire a priori la part des choses entre la partie *synoptique* et la partie *locale* des précipitations : il n'est donc pas possible d'estimer l'incertitude sur la prévision.
- A l'inverse, si l'on se passe de toute exigence au niveau de l'analogie, jusqu'à considérer que toutes les journées de l'archive sont analogues à S^0 , alors la distribution obtenue correspond à la distribution marginale des précipitations *i.e.* la distribution climatologique. Cette distribution inclut très certainement la situation finalement observée, mais n'est pas vraiment informative en terme de prévision !

Le corollaire de ces considérations est que, pour que les prévisions par analogie aient une quelconque utilité, il faut que la partie des précipitations expliquée par les conditions locales soit plus faible que celle expliquée à travers les variables synoptiques utilisées. Cela forme l'une des hypothèses de la méthode. Au vu des résultats obtenus (cf. supra), il semblerait qu'elle soit effectivement vérifiée. La question qui demeure est : jusqu'où peut-on raffiner les critères d'analogie pour que la variance de l'effet des conditions locales soit la plus faible possible ?

1.3.5 Objectif de la thèse

Jusqu'à présent, l'une des difficultés majeures dans la mise en place d'une telle méthode de prévision résidait dans la nécessité de disposer d'une *bonne* archive météorologique. Les archives constituées ne renfermaient qu'un nombre très limité de variables atmosphériques et rien n'assurait leur homogénéité dans le temps.

Néanmoins, l'apparition récente de nouvelles archives, *les réanalyses* (cf. 2.1.3), semble avoir résolu ce problème. On dispose maintenant ainsi de nombreuses variables météorologiques qui permettent une description plus fine de l'atmosphère. Ces nouveaux outils nous ont donc incité à revisiter la méthode de prévision des précipitations par analogie actuellement utilisée chez EDF. On peut maintenant tenter l'utilisation de nouvelles variables d'analogie afin de mieux décrire l'état atmosphérique et de réduire l'incertitude actuelle des prévisions.

L'objectif du travail présenté ici est donc de tester diverses modifications de la méthode actuelle, définie comme étant la méthode de référence. Nos prédicteurs restent des pluies moyennes sur un bassin versant (ou un groupement pluviométrique, cf. 2.2) cumulées sur 24 heures. Nous nous attacherons cependant à exprimer les prévisions sous la forme de fonction de répartition de précipitation prévue, et non pas sous la forme des trois quantiles actuellement utilisés.

Les différents tests vont être effectués en prévision parfaite, c'est-à-dire que les variables d'analogies seront des variables observées et non pas prévues. De cette manière, nous obtiendrons des résultats indépendants de tout modèle NWP. Il faut néanmoins garder à l'esprit que cette méthode a vocation, in fine, à être utilisée en adaptation des modèles numériques.

2

Les données utilisées

« *Mais voici, blanche et diaphane,
La Mémoire, au bord du chemin,
Qui me remet, comme Ariane,
Son peloton de fil en main.* »

Théophile Gautier

Comme nous l'avons vu dans le chapitre précédent, une recherche d'analogues fait appel à la mémoire des situations passées. Plus celle-ci est riche et précise, meilleure sera la prévision. Il est donc nécessaire de se constituer une mémoire conséquente, sous la forme d'importants jeux de données.

Dans la pratique, ces données sont organisées en deux archives distinctes qui couvrent toutes deux la même période :

- *une archive météorologique* de variables synoptiques qui sert à sélectionner les journées analogues : ces variables forment les *prédicteurs* ;
- *une archive de predictands* qui permet d'associer une valeur de la variable à prévoir à chaque situation reconnue comme analogue. Dans notre cas, la variable cible étant la précipitation, nous utiliserons une archive pluviométrique.

2.1 L'archive météorologique

2.1.1 Contraintes et contradictions

La création d'une archive météorologique est une étape cruciale dans la mise en place d'une méthode de prévision par analogie. Il est effectivement nécessaire de respecter plusieurs contraintes plus ou moins contradictoires. Un compromis doit ainsi être trouvé entre :

- ① **La longueur de l'archive** : plus la période couverte par l'archive est étendue, plus il sera facile de trouver un nombre suffisant d'analogues pour toutes les situations météorologiques que l'on sera amené à rencontrer.

- ② **La richesse de l'information contenue** : l'utilisation de plusieurs variables d'analogie pertinentes devrait permettre d'assurer une meilleure discrimination entre les situations générant ou non de fortes précipitations. Idéalement, toutes les variables intervenant dans la genèse des précipitations devraient y figurer, quitte à ce que certaines soient redondantes. Les algorithmes auront la charge de décider lesquelles sont les plus pertinentes au moment de l'optimisation de la méthode.
- ③ **L'homogénéité de l'archive sur toute la période couverte** : des changements de capteurs de mesure ou d'algorithmes d'analyse au cours de la période d'archivage peuvent entraîner une modification de la texture des données au cours du temps. On peut alors se demander si ces données sont toujours comparables entre elles.

La longueur d'archive nécessaire dépend bien sûr de l'extension spatiale de la fenêtre d'analogie considérée. Une cinquantaine d'années semble suffire pour être en mesure de trouver au moins quelques analogues correctes sur un champ de géopotentiels en considérant une fenêtre réduite, de l'ordre de $20^\circ \times 20^\circ$, alors que si l'on s'intéresse à des domaines plus grands, on atteint vite des longueurs d'archivage astronomiques (Van Den Dool 1994)! En outre, la longueur de l'archive impose la limite de ce qui est prévisible. Si aucune des situations de l'archive n'a engendré de pluie centennale, on ne sera jamais en mesure de prévoir une pluie centennale. Ainsi, une archive de 50 ans contient en moyenne 5 situations décennales, si bien que si l'on sélectionne 50 analogues, il n'y en a que 5 qui peuvent, dans le meilleur des cas, atteindre la valeur de la pluie décennale. Cela implique que tous les quantiles inférieurs au quantile 90% restent systématiquement en deçà de la pluie décennale. La prévision est donc biaisée dans le cas des valeurs extrêmes, du fait de la limitation de l'extension temporelle de l'archive (nous reviendrons plus longuement sur ce point lors de la Partie III).

C'est pourquoi il est nécessaire d'utiliser au moins plusieurs décennies de données. Or, jusqu'à très récemment, si l'on souhaitait disposer d'une archive à la fois longue et homogène, il était nécessaire de se limiter à des variables robustes :

- qui dépendent peu des modèles météorologiques utilisés pour leur analyse, afin que les fréquents changements de modèle opérationnel n'aient pas d'impact majeur sur ces variables,
- et qui ont une échelle spatiale suffisamment grande pour que les évolutions du réseau de mesure aient peu d'influence (le réseau de radiosondage a ainsi été significativement diminué, notamment sur l'atlantique nord, à la suite des crises économiques des années 70).

En outre, pour que l'adaptation par analogie soit utilisée opérationnellement en prévision de manière utile, il faut que les variables d'analogie utilisées puissent être correctement prévues! Si leur prévision est fantaisiste, la recherche des analogues se référera à des situations qui ne se produiront pas : les prévisions de précipitation seront donc elles aussi fantaisistes.

Du fait de ces contraintes, il n'était pas très réaliste d'utiliser autre chose que des archives de géopotentiels, et parfois, de températures. Mais ces variables sont-elles suffisantes pour définir des analogies pertinentes pour la prévision des précipitations? Ne faudrait-il pas faire appel à d'autres variables, qui apportent notamment des informations sur l'humidité de l'atmosphère?

2.1.2 L'archive opérationnelle EDF

Utilisant de façon opérationnelle une méthode d'adaptation par analogie depuis 1970, la Division Technique Générale de EDF s'est dotée d'une archive météorologique qu'elle a ensuite constamment mise à jour. Nous l'appellerons par la suite *archive EDF*.

Les variables archivées ont été sélectionnées en tenant compte des différentes contraintes exposées précédemment ainsi que de leur caractère a priori informatif vis-à-vis des cumuls de précipitation journaliers :

- le géopotentiel à 700 hPa, a priori sensible aux grandes lignes de la circulation atmosphérique,
- le géopotentiel à 1000 hPa, a priori plus sensible aux conditions locales des écoulements atmosphériques,
- et la température à 700 hPa, qui n'est plus utilisée aujourd'hui.

Ces variables sont archivées chaque jour à 0 hTU et l'archive EDF couvre maintenant la période 1953-2002.

La constitution de cette archive a été initiée au début des années 70 lors de la mise en place de l'outil de prévision. Les modèles météorologiques numériques en étaient encore à leurs débuts à cette époque, et c'est pourquoi cette archive ne contenait que des variables brutes, c'est-à-dire directement mesurées. Elle était donc constituée par des mesures issues du réseau alors existant de 37 stations de radiosondage couvrant l'Europe Centrale, l'Europe de l'Ouest et le proche Atlantique (figure 2.1). Ce réseau n'est pas resté stable au cours du temps et sa densité est allée

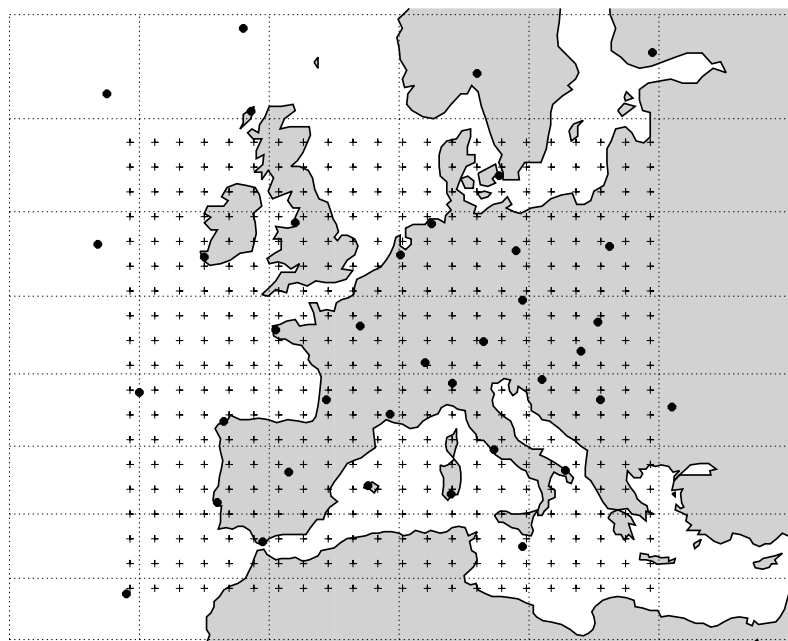


FIG. 2.1: Emplacement des 37 stations de radiosondage (ronds noirs) et des nœuds de la grille utilisée par l'archive EDF (croix).

en diminuant. Les stations manquantes ont alors dues être reconstituées à partir des stations proches restées disponibles.

Au milieu des années 80, devant la diminution du nombre de stations de radiosondage et les progrès des modèles numériques, Météo-France, fournisseur des données, a préféré fournir non plus les mesures brutes mais plutôt les valeurs analysées de ces variables. Celles-ci étaient ensuite interpolées au niveau de chacune des 37 stations de radiosondage initiales. Ainsi, les données contenues dans l'archive sont devenues dépendantes des changements de modèles qui s'opéraient à Météo-France, mais leur qualité profitait de l'apport d'une analyse météorologique.

A la suite des travaux de Guilbaud (1997), les données de l'archive ont été exprimées sur une grille régulière en vue de l'utilisation du critère d'analogie S1 (cf. 4.1.1.2 pour sa définition). Pour ce faire, les données ont été interpolées par une méthode *spline* de type *plaque mince*, sans qu'il n'y ait de contraintes liées aux équations qui régissent les mouvements atmosphériques. La grille choisie est une grille régulière selon une projection Mercator, dont les mailles font à peu près, au niveau de la France, $2^\circ \times 1.5^\circ$ (figure 2.1).

Pour ces raisons historiques, l'archive EDF ne présente pas de grandes garanties d'homogénéité et ne contient qu'un nombre restreint de variables. Mais elle a l'immense mérite d'exister et d'être toujours maintenue, ce qui n'est pas négligeable !

2.1.3 Les Réanalyses

L'archive EDF n'est donc pas parfaite, mais jusqu'à très récemment, c'était la seule existante. Maintenant, un nouveau type d'archive est apparu : les réanalyses.

2.1.3.1 Qu'est-ce qu'une réanalyse ?

Face aux interrogations portant sur d'éventuels changements climatiques, la communauté des climatologues a exprimé le besoin de disposer d'archives fiables du climat actuel. En effet, les nombreuses ruptures présentes dans les séries de données « *classiques* » sont essentiellement dues à des changements de position d'un capteur ou à un changement dans l'analyse de la variable considérée. Or, étant du même ordre de grandeur que celui attendu actuellement d'un changement climatique, ces ruptures artificielles masquent les évolutions réelles dans les séries de données.

De grands centres météorologiques ont alors décidé de mettre sur pied de grands projets de réanalyses. Ceux-ci consistent, dans un premier temps, à rassembler toutes les mesures atmosphériques disponibles au cours des dernières décennies (mesures au sol, radiosondages, images satellites, ...). Puis, dans un deuxième temps, à les analyser toutes ensemble avec un modèle météorologique unique sur toute la période de réanalyse.

Cette analyse permet d'interpoler ces données en respectant les contraintes dues aux lois physiques qui régissent les phénomènes atmosphériques, contrairement à une simple interpolation '*mathématique*' comme celle utilisée pour l'archive EDF. Etant effectuée a posteriori, l'analyse prend en compte non seulement la cohérence spatiale des données, mais aussi leur cohérence

temporelle. Elle intègre en outre des données qui ne sont habituellement pas encore disponibles en temps réel. De cette manière, les mesures aberrantes peuvent être corrigées, les données manquantes reconstituées, et l'interpolation en point de grille effectuée de manière satisfaisante même dans des régions relativement pauvres en données mesurées.

Evidemment, même si elles représentent un net progrès par rapport à ce qui existait, ces archives elles non plus, ne sont pas parfaites. D'une part, les réseaux de mesure ont évolué : les réseaux 'classiques' ont eu tendance à voir leur densité diminuer, alors que sont apparus de nouveaux capteurs tels que les satellites, lesquels apportent de nouveaux types de données. Les mesures sur lesquelles s'appuient les réanalyses ne sont donc pas homogènes dans le temps. Ainsi, pour une région donnée du globe, la qualité des réanalyses est variable dans le temps, en fonction de l'évolution des réseaux de mesure, même si l'utilisation d'un modèle météorologique minimise cette variabilité. Les réanalyses sont donc aussi homogènes *que possible*.

D'autre part, il convient de différencier les variables qui dépendent peu du modèle utilisé pour procéder à l'analyse, de celles qui en dépendent fortement. Ainsi, les variables liées à la partie dynamique du modèle, telles que les géopotentiels, sont quasiment indépendantes de celui-ci et sont plus ou moins imposées par les mesures disponibles. A l'inverse, les variables reliées à la partie physique du modèle, telles que l'humidité par exemple, peuvent varier grandement d'un modèle à l'autre. Ce dernier type de variable, même s'il reste homogène tout au long d'une archive de réanalyse, puisque le modèle utilisé ne varie pas, ne l'est pas forcément avec le modèle utilisé en temps réel pour analyser un jour courant, ou même avec la réalité.

2.1.3.2 Les Réanalyses NCEP/NCAR

Le NCEP (National Center for Environmental Prediction, anciennement NMC) et le NCAR (National Center for Atmospheric Research) se sont associés dès 1991 afin de produire 40 années de réanalyses (de 1957 à 1996). Le projet s'est depuis étendu et les réanalyses NCEP/NCAR sont maintenant disponibles de 1948 jusqu'au mois précédent de l'année en cours. Le projet ne s'est en effet pas arrêté à 1996. Il continue toujours à analyser de la même façon les nouvelles observations, afin que les événements actuels puissent être comparés avec les événements passés.

Ces réanalyses sont produites avec un modèle météorologique correspondant à l'état de l'art des modèles opérationnels tel qu'il était en 1995 (Kalnay *et al.* 1996). C'est un modèle spectral global ayant une troncature horizontale T62 (≈ 210 km) et 28 niveaux verticaux.

Les variables disponibles sont extrêmement nombreuses et couvrent tous les aspects des processus atmosphériques, depuis les variables dynamiques jusqu'à l'ensemble des éléments du bilan radiatif. Elles ont été classées en 3 catégories, selon leur degré de dépendance au modèle utilisé :

- **type A** : variables les plus robustes, dépendantes essentiellement des mesures (pression, température ou vent dans les hautes couches de l'atmosphère par exemple),
- **type B** : variables dépendantes d'observations mais très influencées par le modèle (humidité ou température de surface par exemple),
- **type C** : variables non mesurées et dépendantes exclusivement du modèle (tourbillon potentiel, nébulosité ou précipitation par exemple).

Code	Variable	Type	Niveaux (hPa)							heure (hTU)	
HGT	Altitude du géopotiel	A	1000	850	700	500	300	200		00	12
TMP	Température	A		850		500		200		00	12
UGR	Vitesse du vent zonal	A		850		500		200		00	12
VGR	Vitesse du vent méridien	A		850		500		200		00	12
RHU	Humidité relative	B		850		500				00	12
VVE	ω (<i>i.e.</i> vitesse verticale en coordonnées « <i>presion</i> »)	B				500				00	12
PWA	Eau précipitable	B	cumul sur la colonne d'air							00	12

TAB. 2.1: Variables de l'archive LTHE définies selon des niveaux géopotentiels.

Code	Variable	Type	Niveaux (° K)		heure(hTU)
UG2	Vitesse du vent zonal	A	315	330	00
VG2	Vitesse du vent méridien	A	315	330	00
TM2	Température	A	315	330	00
PVO	Tourbillon potentiel	B	315	330	00

TAB. 2.2: Variables de l'archive LTHE définies selon des surfaces isentropiques.

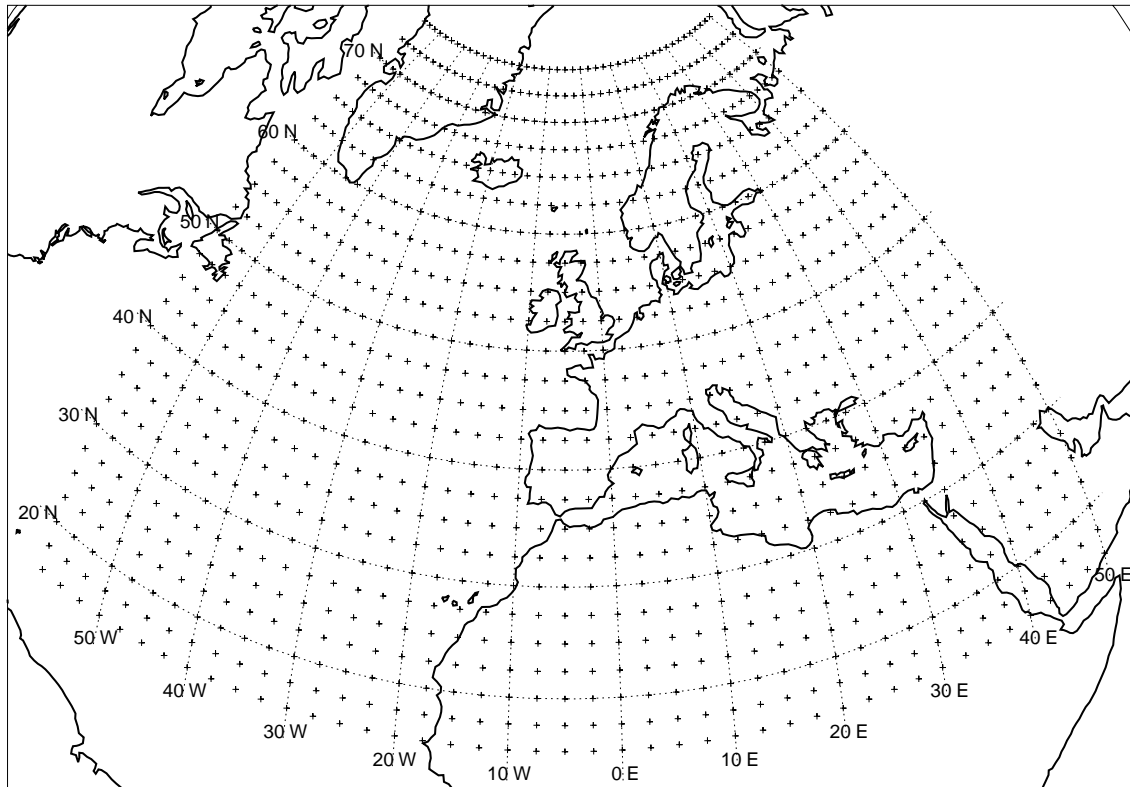


FIG. 2.2: Domaine d'extraction et grille de stockage de l'archive LTHE.

Nous avons utilisé ces réanalyses en nous appuyant sur les CD-Rom annuels édités par le NCAR. Chacun d'entre eux couvre une année entière de réanalyse et contient une sélection de variables définies selon une résolution horizontale de $2.5^\circ \times 2.5^\circ$ sur l'ensemble du globe. Nous nous sommes créés notre propre archive en conservant à notre tour, parmi cette sélection, toutes les variables qui peuvent présenter un intérêt quelconque pour la recherche d'analogie. Celles-ci sont définies soit sur des niveaux géopotentiels, disponibles à 00 et 12 hTU (tableau 2.1), ou soit sur des niveaux isentropiques¹, disponibles seulement à 00 hTU (tableau 2.2). Nous avons par ailleurs limité notre extraction à un domaine allant de 57.5°E à 50°W et de 15°N à 75°N (figure 2.2). Par la suite, nous appellerons cette archive extraite des réanalyses NCEP/NCAR l'*archive LTHE*.

2.1.3.3 Les Réanalyses ERA-15 et ERA-40

Le CEPMMT a mis en place deux projets successifs de réanalyses. Le premier, le projet ERA-15, a pris place au milieu des années 90 et couvre la période 1979-94. Il est contemporain aux réanalyses NCEP/NCAR mais le modèle employé fait d'autres choix en matière de paramétrisation des processus physiques et utilise un modèle global spectral de résolution légèrement supérieure : troncature T106 (≈ 125 km) et 31 niveaux verticaux (Gibson *et al.* 1997). Cependant, la période de réanalyse est trop courte pour nous intéresser ici.

Le projet ERA-40, qui couvre la période 1957-2001, a succédé aux réanalyses ERA-15. Il a vu sa première version s'achever durant l'été 2003 mais certaines parties de l'archive devaient encore être recalculées suite à la découverte d'anomalies. Le modèle utilisé a été amélioré par rapport à celui des réanalyses ERA-15, notamment au niveau de l'introduction de nouvelles mesures et de nouvelles techniques d'assimilation. La résolution a par ailleurs été notablement augmentée : le modèle utilise maintenant une troncature T159 sur 60 niveaux verticaux (Uppala *et al.* 1999).

L'archive ERA-40 présente une résolution spatiale de $1^\circ \times 1^\circ$ et apporte d'intéressantes perspectives en matière de qualité des résultats. Malheureusement, elle n'était pas encore disponible pour être utilisée dans le cadre de cette thèse.

¹Une surface isentropique est une surface à température potentielle (θ) constante. La température potentielle d'une particule d'air correspond à la température de cette particule, supposée sèche, quand elle est ramenée à un niveau de pression de 1000 hPa de façon adiabatique, c'est-à-dire sans échange de chaleur avec l'atmosphère. Cette variable est couramment utilisée en météorologie car elle est un invariant des masses d'air au cours d'évolutions adiabatiques réversibles.

L'altitude d'un niveau isentropique peut très fortement varier selon les conditions atmosphériques du moment. Néanmoins, si l'on se place dans le cadre de l'atmosphère standard, on peut considérer les altitudes suivantes :

$\theta=315^\circ\text{K}$ correspond à une altitude de 7 à 7.5 km (niveau de la surface isobare 400 hPa)

$\theta=330^\circ\text{K}$ à une altitude de 11 km (niveau de la tropopause)

Le lecteur désirant approfondir ces notions est invité à consulter l'ouvrage de Triplet et Roche (1986).

2.2 L'archive pluviométrique

L'archive pluviométrique contient les mesures historiques de précipitation sur les différents bassins versants qui vont intéresser la prévision.

2.2.1 Du cumul des précipitations

La pluie n'étant qu'une succession de chutes de gouttes d'eau distinctes les unes des autres, une valeur de pluie ponctuelle et instantanée n'existe pas. Les précipitations sont donc toujours *cumulées sur une durée et moyennées sur une surface*.

Pour constituer une archive pluviométrique, il faut donc préalablement choisir une superficie et une durée de cumul. Bien que plus ou moins arbitraires, ces choix doivent être compatibles avec :

- les besoins de l'utilisateur de la prévision, puisqu'ils imposent la résolution temporelle et spatiale des prévisions,
- et l'archive météorologique dont on dispose, étant donné que les cumuls de précipitations doivent être explicables par les données météorologiques. En effet, si les données contenues dans l'archive météorologique ne sont définies que toutes les 24 heures, il paraît par exemple difficile de prévoir des cumuls de précipitations infra journaliers.

En ce qui nous concerne, nous disposons de données météorologiques toutes les 12 heures au sein de l'archive LTHE (00 et 12 hTU). Nous pourrions donc envisager de considérer des précipitations cumulées sur 12 heures. Malheureusement, en France, les longues archives de précipitations disponibles à un pas de temps infra journalier sont peu nombreuses comparées aux archives journalières. C'est pourquoi, afin de disposer d'un panel de données suffisamment fourni, nous nous sommes limités à un pas de temps journalier. Il est important de noter que, du fait des protocoles pratiques des mesures pluviométriques, la mesure associée au jour J correspond en fait au cumul mesuré entre 6 hTU le jour J et 6 hTU le jour J+1 (et non pas entre 00 et 24 hTU le jour J).

Le choix de la superficie sur laquelle sont moyennés les cumuls se fait en liaison avec celui de la durée du cumul. Idéalement, il convient de choisir une superficie qui regroupe une zone pluviométriquement homogène pour la durée utilisée. Si l'on considère ainsi une surface trop grande, on est amené à mélanger différentes influences. L'analogie se doit alors d'être moins spécifique, et par là même, moins performante. De plus, la prévision porte sur une variable hétérogène qui présente moins d'intérêts pour l'utilisateur : les précipitations ne concernent souvent qu'une partie de la surface considérée.

A titre d'exemple, considérons les pluies journalières sur les bassins de la Loire et de l'Allier. Bontron *et al.* (1999) ont montré, sur un échantillon de 108 pluviomètres, qu'entre deux points situés à 25 km l'un de l'autre, ces précipitations présentaient un coefficient de corrélation de 0.5 et qu'à partir de 100 km, elles étaient décorréliées (figure 2.3). Cette distance de 100 km correspond à la *portée* des pluies journalières dans cette région. On peut donc considérer que la superficie

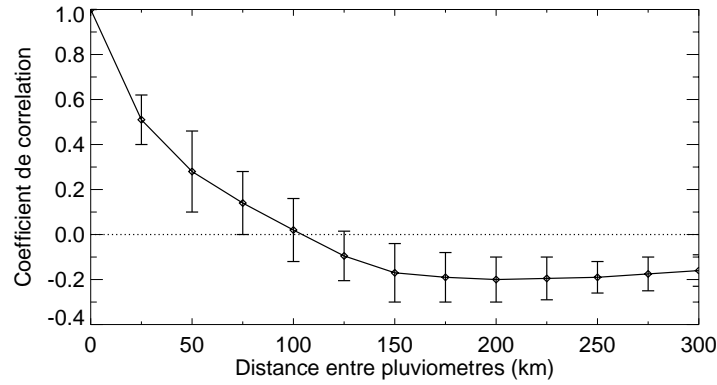


FIG. 2.3: Corrélogramme de la pluie journalière sur 108 stations des bassins de la Loire et de l'Allier (d'après Bontron et al. 1999).

requis ici pour obtenir des précipitations journalières homogènes correspond à une surface de 25 km de rayon (un quart de la portée), c'est à dire une superficie de 1900 km².

D'une façon plus générale, Lebel *et al.* (1987) ont établi que, dans la région des Cévennes, la portée d (en km) s'exprime en fonction de la durée de cumul Δt (en heures) de la façon suivante :

$$d \approx 25 \cdot \Delta t^{0.3} \quad (2.1)$$

Pour une même durée de cumul Δt de 24 heures, la portée est donc plus faible dans les Cévennes, où $d = 65$ km (soit une superficie de seulement 830 km²), que dans le bassin de la Loire. Cette équation nous permet d'estimer, au moins pour cette région, les superficies associées à des durées de cumul plus petites : soit 450 km² pour 12 heures et 350 km² pour 6 heures.

L'existence de phénomènes localisés nous encourage par ailleurs à considérer des superficies qui ne soient pas trop petites non plus. En effet, les processus qui déterminent précisément l'endroit où un orage prendra naissance par exemple, se déroulent à des échelles largement inférieures à celles disponibles dans l'archive météorologique. On conçoit aisément qu'il n'est pas forcément possible de les localiser a priori avec exactitude et que l'on augmente nos chances de prévision en considérant une surface plus importante, au risque de « diluer » l'orage en le moyennant sur une surface trop grande.

Parallèlement à la question de la superficie sur laquelle sont moyennées les précipitations, se pose la question de son échantillonnage. En effet, la valeur estimée est fonction de quelques mesures quasi ponctuelles, localisées aux postes pluviométriques, et intégrées par une simple moyenne arithmétique. Pour avoir une précision suffisante, il est donc intéressant d'avoir un nombre suffisant de postes de mesure, supérieur à ce qu'indique la portée moyenne théorique des précipitations :

- d'une part pour prendre en compte d'éventuels effets de dérive (différences de pluviométrie entre les piedmonts et les crêtes par exemple),
- et d'autre part pour éviter que des phénomènes localisés, comme des cellules orageuses estivales, ne passent entre les points de mesure.

D'une manière générale, on considère que les mesures par pluviomètres ont une précision ponctuelle moyenne de l'ordre de 10%, la principale cause d'incertitude étant l'effet du vent. Lorsque l'on considère plusieurs mesures voisines, l'erreur sur la moyenne est moindre à condition que l'échantillonnage spatial soit suffisant. Dans ce cas, si les erreurs de mesure sont indépendantes, l'erreur moyenne diminue comme l'inverse de la racine carrée du nombre de postes : ainsi, si l'on dispose de 4 postes, elle est divisée par 2. Pour les bassins qui nous concernent, l'ordre de grandeur de l'incertitude peut donc être estimé à quelques pourcents ($\approx 5\%$).

2.2.2 Les groupements pluviométriques

Nous ne travaillons pas, au sens strict du terme, sur des bassins versants, mais plutôt sur des *groupements pluviométriques*. Plusieurs pluviomètres, pris au sein d'une zone pluviométriquement homogène, servent ainsi à définir une lame d'eau moyenne même s'il n'appartiennent pas au même bassin hydrographique. Les groupements que nous utilisons sont en outre constitués de telle sorte que la densité de pluviomètres utilisée permet d'approcher la moyenne spatiale par une simple moyenne arithmétique des postes pluviométriques.

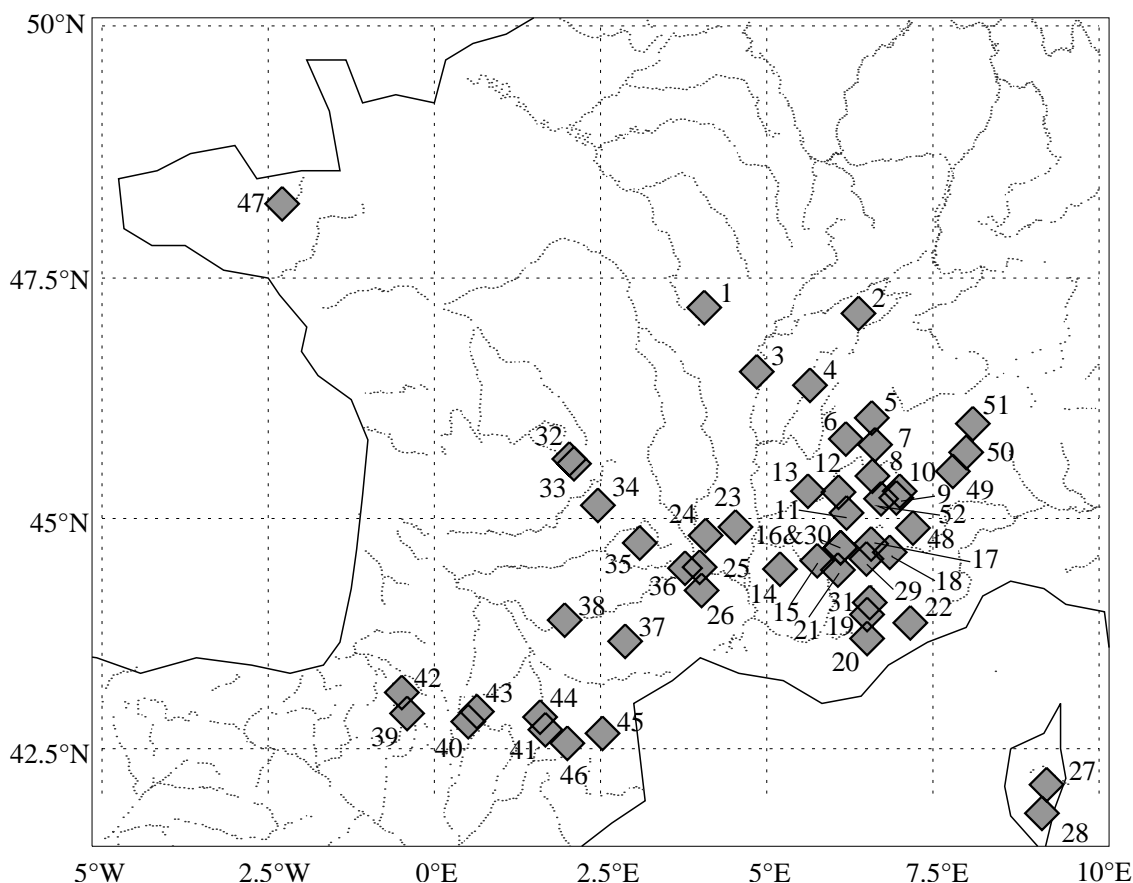


FIG. 2.4: Carte des groupements pluviométriques EDF.

n°	Groupelement	F0 (%)	R10 (mm)	Nombre de postes
1	Cure	46	54	4
2	Doubs	42	67	4
3	Saône	52	65	3
4	Ain-Valserine	45	83	6
5	Arve-Giffre	49	71	6
6	Fier	52	81	4
7	Arly	52	75	3
8	Hte_Isère-Doron	55	63	5
9	Piedmont_italien	54	90	4
10	Hte_Maurienne	55	92	2
11	Romanche-Drac_Inf.	56	61	5
12	Isère_moyenne	51	76	3
13	Vercors-Chartreuse-Nd._Isère	49	83	3
14	Drôme	65	78	3
15	Buëch	60	81	3
16	Drac	58	95	5
17	Hte_Durance	60	66	5
18	Guil_Ubaye	63	65	3
19	Verdon	61	77	4
20	BVI_Verdon	63	94	4
21	Durance_moyenne	64	79	4
22	Var-tinée-Roya	58	101	7
23	Doux-Eyrieux	56	137	4
24	Loire_supérieure	46	170	5
25	Chassezac	56	239	3
26	Gard-Cèze	54	154	5
27	Corse_orientale	63	128	2
28	Corse_du_sud	66	113	2
29	MORDOR_Durance	57	67	6
30	MORDOR_Drac	58	90	5
31	MORDOR_Verdon	63	81	3
32	Vézère-Vienne-Dordogne	44	68	6
33	Creuse-Cher-Sioule	40	50	10
34	Cere-Marognes-Rhue-Dordogne_RG	41	64	13
35	Bromme-Bes-Borald.-Truy.	44	65	10
36	Ht_Tarn-Ht_Lot-Ht_Allier	41	102	19
37	Ht_Agout-Arn-Dourdou	48	129	9
38	Aveyron-tarn_Inf.-Garonne_Inf.	47	67	8
39	Pyrénées_Atl._Ht_bassin	45	72	4
40	Pyrénées_Cent._Ht_bassin	47	75	4
41	Pyrénées_Arieg_Ht_bassin	47	86	6
42	Pyrénées_Atl._Piedmont	42	69	7
43	Pyrénées_Cent._Piedmont	41	77	4
44	Pyrénées_Arieg_Piedmont	50	81	4
45	Tech	64	123	2
46	Aude-Tet	50	77	5
47	Bretagne	45	45	4
48	Dora_Riparia	52	74	9
49	Sésia	48	156	9
50	Toce_globale	49	144	9
51	Toce_Nord	51	136	6
52	Arc_à_Bramans	51	75	5

TAB. 2.3: Caractéristiques des groupements pluviométriques EDF (F0 : fréquence des valeurs nulles ; R10 : valeur de pluie journalière décennale).

Nous avons utilisé deux séries de groupements qui couvrent toutes deux la période 1953-2001 :

- **Les groupements EDF** : ces groupements nous ont été aimablement fournis par la Division Technique Générale de EDF. Ce sont les groupements sur lesquels portent leurs prévisions opérationnelles, auxquels a été adjoint un groupement concernant le bassin de l'Arc à Bramans. Ce sont donc des groupements opérationnels qui traduisent les besoins d'un utilisateur tel que EDF.
- **Les groupements RIO** : ces groupements ont été financés par l'appel d'offre RIO II du Ministère de l'Environnement et du Développement Durable. Une partie de ces bassins concerne là encore des besoins exprimés (groupements de Marseille, de la Siagne, du Paillon, ...), mais la plupart ont été créés afin de tester la capacité de résolution spatiale de l'adaptation par analogie. Ces groupements, disponibles plus tardivement que les groupements EDF, ne seront pas utilisés lors de l'optimisation de la méthode (partie II).

Ces groupements sont, pour une grande part, situés dans le quart Sud-Est de la France (figure 2.4 et 2.5). Ils présentent des caractéristiques pluviométriques (valeurs de pluies décennales et fréquence des valeurs nulles, tableaux 2.3 et 2.4) ainsi que des superficies relativement variées, qui vont de 500 à 2000 km². Les superficies associées à chaque groupement sont cependant très approximatives, ces groupements ne reposant pas forcément sur des limites hydrographiques strictes, comme on l'a signalé. Cela explique pourquoi certains portent le même nom dans les deux archives : ils correspondent à deux échantillonnages différents du réseau pluviométrique d'une région. Il est donc normal qu'ils présentent des caractéristiques différentes.

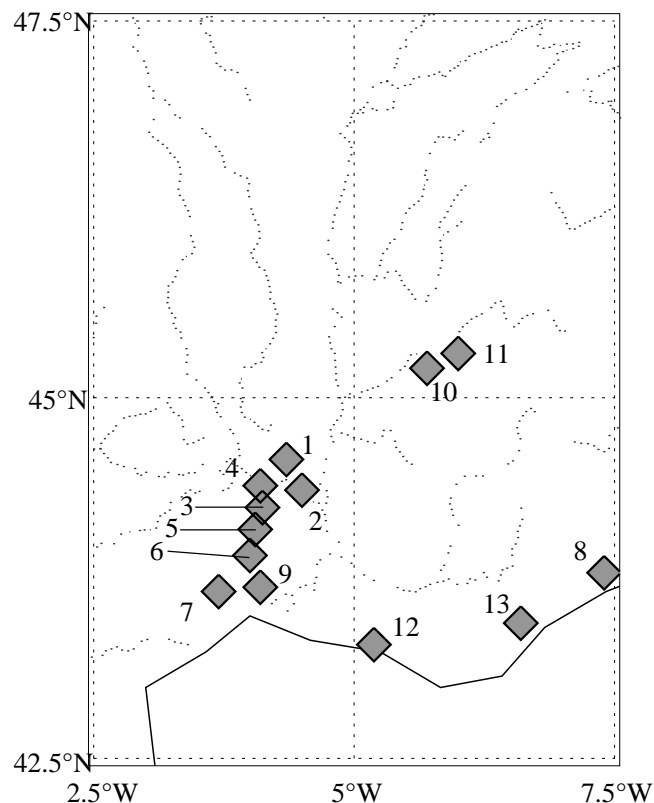


FIG. 2.5: Carte des groupements pluviométriques RIO.

n°	Groupelement	$F0$ (%)	$R10$ (mm)	Nombre de postes
1	Haute_Ardèche	61	180	6
2	Basse_Ardèche	65	121	5
3	Ceze	64	153	4
4	Chassezac	55	171	5
5	Gardon_Alès	59	148	4
6	Gardon_Anduze	61	163	6
7	Herauld	51	202	4
8	Paillon	66	111	8
9	Vidourle	67	129	5
10	Grenoble	51	72	8
11	Belledonne_Grésivaudan	49	62	14
12	Grand Marseille	73	84	7
13	Siagne	66	99	9

TAB. 2.4: Caractéristiques des groupements pluviométriques RIO ($F0$: fréquence des valeurs nulles ; $R10$: valeur de pluie journalière décennale).

2.2.3 Transformation des précipitations

Nous n'allons pas travailler directement avec les valeurs brutes de précipitations. La pluie est en effet une variable assez particulière qui mérite quelques traitements avant utilisation. Nous lui faisons donc subir deux transformations :

- ① Nos différents groupements, soumis à des régimes pluviométriques variés, sont difficilement comparables entre eux. Ainsi, 100 mm de pluie en une journée n'a pas le même impact sur un groupement tel que **Bretagne**, dont la pluie décennale est de moins de 50 mm, ou sur un autre, tel **Chassezac**, qui connaît pour sa part une pluie décennale de plus de 200 mm. Nous appliquons donc un facteur d'échelle à tous les groupements, afin qu'une même valeur ait une signification comparable pour tous. Pour ce faire, nous avons décidé d'exprimer les précipitations en terme de fraction de la pluie décennale $R10$ du groupement considéré. De cette manière, les 100 mm de pluie de l'exemple précédent correspondent maintenant à $2.2 R10$ sur **Bretagne**, alors qu'elle ne représente plus que $0.42 R10$ sur **Chassezac**.
- ② Par ailleurs, la pluie journalière est une variable très dissymétrique. Elle est bornée en 0, où elle présente un point d'accumulation, alors qu'elle n'est pas bornée sur les grandes valeurs. A titre d'exemple, la figure 2.6a présente la fonction de répartition de la pluie journalière sur le groupement **Doubs**. Comme le montre cette figure, une surestimation de 5 mm sur une précipitation de 3 mm est autrement plus conséquente en terme d'écart en fréquence d'apparition que la même erreur sur une pluie de 20 mm. Or, lorsque l'on va chercher à estimer les performances d'un système de prévision, nous allons être amenés à calculer des écarts de précipitations, entre valeurs observées et valeurs prévues. C'est pourquoi, afin de donner moins de poids aux écarts sur les fortes valeurs, il est courant de considérer la racine carrée des précipitations.

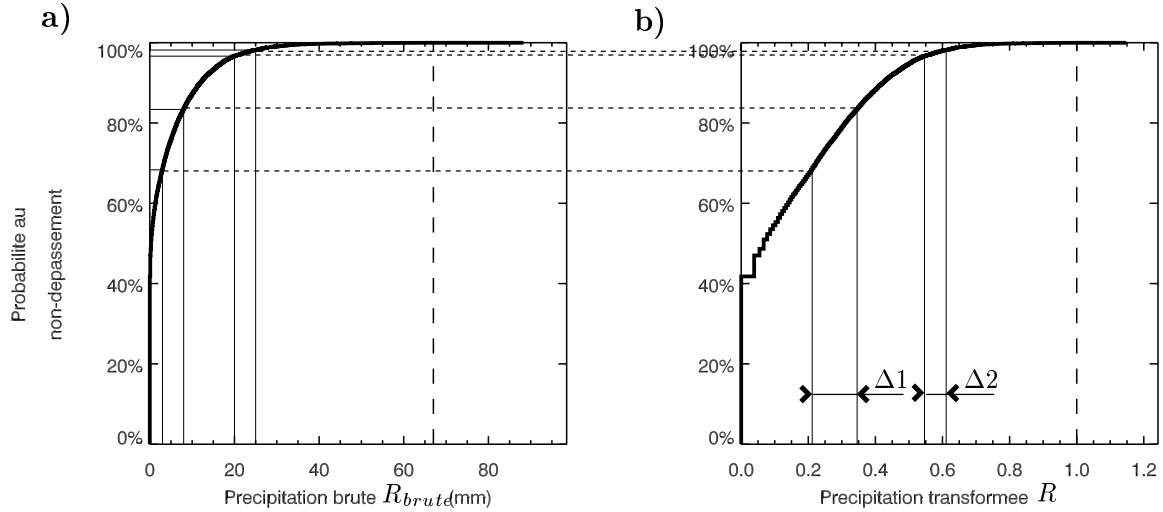


FIG. 2.6: Fonction de répartition du groupement Doubs établie sur la période 1953-96 (pluie au pas de temps journalier). **a)** : pluie brute; **b)** pluie transformée.

Finalement, nous allons travailler avec la variable R qui correspond à la transformation des précipitations brutes R_{brutes} selon l'équation suivante :

$$R = \sqrt{\frac{R_{brutes}}{R10}} \quad (2.2)$$

La pluie décennale $R10$ est calculée par un simple ajustement de la loi de Gumbel sur les maxima annuels (période 1953-96). Ce calcul rapide, même s'il n'est pas suffisamment précis pour effectuer un dimensionnement d'ouvrage, nous fournit un facteur d'échelle du comportement de chaque groupement tout à fait adapté à nos besoins.

La figure 2.6b montre l'effet de cette transformation sur la distribution empirique des précipitations. Exprimé en variable transformée, l'écart de 5 mm sur une observation de 3 mm ($\Delta 1$ sur la figure) est maintenant considéré comme deux fois plus important que quand il s'applique à une observation de 20 mm ($\Delta 2$ sur la figure).

2.3 Format d'archivage

L'ensemble des archives utilisées constitue une masse de données relativement conséquente. En pratique, son stockage n'est pas évident. Afin de conserver des fichiers d'archives à la fois de taille compacte et rapides d'accès, nous avons choisi d'utiliser le format NetCDF (Rew et Davis 1997).

Comme tout format de stockage standardisé, il impose certaines contraintes au moment de la création des fichiers de données. Mais, une fois cette tâche accomplie, il offre de nombreux avantages :

- Les fichiers NetCDF sont portables d'un système d'exploitation à un autre, quelle que soit la façon dont ces systèmes codent dans leurs mémoires un entier ou un réel en virgule flottante. C'est à dire qu'une archive créée par exemple sous Unix pourra sans aucun problème être lue sous MS-DOS.
- Il permet un accès direct à n'importe quelle donnée située n'importe où dans l'archive, sans avoir à passer en revue l'intégralité du fichier.
- Ce format est auto-explicatif, c'est à dire que toutes les informations nécessaires à la compréhension de ce que contient un fichier sont incluses dedans. Ainsi, il est possible de prêter un fichier NetCDF à un autre utilisateur sans avoir à expliquer comment avoir accès aux données. Cela évite de se retrouver au bout de quelque temps l'heureux propriétaire d'un superbe fichier que plus personne (y compris soi même) ne sait comment lire...

Une documentation détaillée mais parfaitement abordable est disponible en ligne à l'adresse suivante : <http://www.unidata.ucar.edu/packages/netcdf/>

3 Qualité et Utilité d'une prévision

人皆知有用之用
而莫知无用之用也

« Tout le monde connaît l'utilité de l'utile,
mais personne ne sait l'utilité de l'inutile. »
Zhuangzi, philosophe chinois, 369-286 av J-C

Emettre une prévision, c'est déjà bien. Mais émettre une bonne prévision, c'est mieux !

Fort de cette constatation de bon sens, on se retrouve néanmoins vite confronté à un problème épineux : comment reconnaître une bonne prévision ? Voire : comment reconnaître qu'une prévision est meilleure qu'une autre ?

Une prévision, émise *a priori*, a vocation à être confrontée *a posteriori* à la réalité. C'est par le biais de cette comparaison que l'on peut juger de sa qualité. Mais une prévision peut être présentée de différentes manières comme nous l'avons vu dans la partie 1.1.3. Or, toutes ces formes de pronostics ne pourront évidemment pas être jugées de la même façon. En outre, les objectifs qui poussent à évaluer la qualité d'une prévision peuvent être très variés (voir par exemple Murphy et Daan 1985). De ces objectifs naissent des attentes différentes vis-à-vis des prévisions, et donc diverses définitions de leur qualité.

C'est pourquoi de nombreux critères d'évaluation de cette qualité ont été définis. Malheureusement, aucun ne s'est révélé universel. Le choix de l'un ou l'autre s'effectue donc à la fois en fonction de la forme de la prévision et des objectifs qui ont présidé à sa mise en œuvre. Comme nous le verrons, il est même souvent difficile de se contenter d'un seul critère.

Après avoir expliqué les principes de base utilisés pour mesurer la qualité des prévisions météorologiques, ce chapitre tentera, sans prétendre à l'exhaustivité, de dresser un panorama des différentes méthodes de vérification. Ici, nous nous focalisons exclusivement sur les prévisions quantitatives¹ de variables scalaires (ponctuelles ou en moyennes surfaciques). Les méthodes

¹Le lecteur intéressé par les spécificités des prévisions qualitatives est invité à se reporter par exemple à l'article de Murphy et Brown (1983).

adaptées aux prévisions de champs (et à l'évaluation d'une spatialisation correcte au sein du champ), ne seront pas abordées, même si certaines techniques présentées pour des séries temporelles ponctuelles peuvent aussi être employées dans ce cas (*RMSE* et corrélation par exemple).

Jusqu'à la partie 3.3.3 incluse, nous présentons des méthodes d'évaluation traditionnelles. Ce premier ensemble forme le contexte que l'on juge nécessaire afin d'appréhender la suite de notre travail (Parties II et III). Ensuite, le lecteur pressé pourra passer directement à la conclusion de ce chapitre (3.4), où nous verrons que l'on va tout particulièrement privilégier le score *CRPS* (cf. 3.3.3.3).

Les parties 3.3.4 et 3.3.5 nous semblent originales mais aussi un peu plus ardues... Elles ne sont pas absolument indispensables à la compréhension du lecteur pressé. La partie 3.3.5 propose un cadre plus général à la définition du *CRPS* et justifie son utilisation ici en explicitant ses diverses interprétations : économiques, analytiques et graphiques. L'autre partie (3.3.4), qui n'enthousiasmera que les spécialistes, traite d'un essai d'utilisation de la notion de vraisemblance comme score de vérification, qui s'est malheureusement révélé infructueux dans notre contexte.

3.1 A quoi reconnaît-on une *bonne* prévision ?

3.1.1 Qualité et utilité : une simple différence sémantique ?

Lorsque l'on veut vérifier les performances d'un système de prévision, il est possible d'adopter deux points de vue différents : celui du prévisionniste, et celui de l'utilisateur. Ces deux catégories d'intervenants n'ont pas les mêmes attentes vis-à-vis des prévisions, et cela conduit à définir deux concepts différents (Murphy et Daan 1985) :

- **la qualité** : c'est le point de vue du prévisionniste, qui s'attache à la comparaison entre les valeurs prévues et les valeurs observées.
- **l'utilité** : c'est le point de vue de l'utilisateur, qui s'intéresse au gain ou à la perte (par exemple sous une forme financière) subis lorsque ses décisions sont basées sur la prévision.

Ces deux notions ne sont pas équivalentes. Un gain de qualité pour un système de prévision n'implique pas une augmentation de son utilité, de même que le système qui présente la moins bonne qualité n'est pas forcément le moins utile (Krzysztofowicz 1992). Ces notions sont cependant souvent confondues et il est fréquent que des mesures d'utilité soient présentées comme des mesures de qualité, et vice-versa.

Or, l'utilité d'une prévision en elle-même n'existe pas : elle ne peut se définir que par rapport à un certain utilisateur. Dès lors que l'on parle d'utilité, on fait référence à l'utilisation qui va être faite de la prévision. En effet, la même prévision n'aura pas nécessairement la même utilité pour deux utilisateurs différents. Considérons ainsi l'exemple d'une prévision probabiliste de précipitations journalières par catégories, telle que proposée par le tableau 3.1, alors que le cumul de précipitations observées pour cette journée a été de 47 mm.

- Si l'utilisateur de cette prévision veut savoir s'il est judicieux de répandre un traitement antiparasite sur ses cultures, cette prévision s'avère tout à fait utile. Elle prévoit clairement

	Classes de précipitation R (mm)						
	R = 0	0 < R ≤ 2	2 < R ≤ 5	5 < R ≤ 10	10 < R ≤ 20	20 < R ≤ 50	50 < R
Probabilité d'appartenance à la classe (%) :	3	7	20	40	15	10	5

TAB. 3.1: Exemple de prévision probabiliste de précipitation par classe de pluie.

qu'il va pleuvoir de manière significative, et que si le traitement est effectué, il sera lessivé et à refaire.

- Si l'utilisateur a en charge la gestion d'un barrage de production d'énergie, cette prévision se révèle peu utile. Les quantités de précipitations prévues sont très largement inférieures aux quantités observées : les apports réels ont été bien plus importants que les apports prévus.

Au final, les concepts de qualité et d'utilité conduisent à exprimer des scores de vérification qui s'utilisent de façon analogue. Il convient toutefois, au moment de faire usage d'un score de vérification, de se demander si ce score mesure la qualité ou l'utilité, et, dans ce dernier cas, s'il est légitime de s'identifier à l'utilisateur pour lequel a été défini ce score.

3.1.2 Généralités méthodologiques

3.1.2.1 Pourquoi un score de vérification ?

Historiquement, les scores de vérification sont apparus dans les années 1940-50. Les prévisionnistes ne disposaient alors pas de modèles de prévision numériques. Ils émettaient donc des prévisions probabilistes subjectives (Sanders 1963). L'objectif des scores de vérification n'était alors pas tant de comparer les prévisionnistes entre eux, mais plutôt de les aider dans la formulation de leurs prévisions subjectives. Ces scores devaient en effet encourager l'émission de prévisions « *honnêtes* », *i.e.* qui reflètent l'opinion « *vraie* » des prévisionnistes, et non pas des prévisions biaisées qui leur permettraient d'augmenter artificiellement leur espérance de réussite (Matheson et Winkler 1976)...

En ce qui nous concerne, nos préoccupations sont légèrement différentes. Dans le cadre de la recherche d'améliorations de la méthode de prévision par analogues, nous sommes amenés à comparer plusieurs algorithmes, différents entre eux. Ces algorithmes produisent, à partir des mêmes jeux de données, des prévisions différentes. La question qui se pose alors est de savoir quel est le meilleur algorithme *i.e.* lequel a produit les meilleures prévisions.

Nous avons donc besoin d'une fonction qui, à chaque couple *série de prévisions/série d'observations*, associe une mesure de la qualité ou de l'utilité de la prévision : c'est le Score Global

S . L'algorithme auquel est associé le plus petit Score Global est celui qui est considéré comme étant le plus performant¹.

Habituellement, pour simplifier le problème de la définition de ce score S , on se ramène à la mesure de la performance de chaque prévision considérée individuellement. On définit ainsi un Score Individuel s , qui permet au sein de chaque couple de la *série de prévisions/série d'observations* d'associer une série (s_i) de Scores Individuels. Grâce à une fonction de composition h , on calcule alors S à partir de cette série (s_i) :

$$S^{M1} = S((x_i^{M1}); (x_i^0)) = h(s(x_1^{M1}; x_1^0); s(x_2^{M1}; x_2^0); \dots; s(x_N^{M1}; x_N^0)) \quad (3.1)$$

où :

- N est le nombre d'éléments de la série de prévisions,
- (x_i^{M1}) la série des prévisions obtenues avec la méthode de prévision (ou l'algorithme) $M1$,
- (x_i^0) la série des observations associées,
- S^{M1} le Score Global de la méthode $M1$.

Les observations (x_i^0) , qui forment notre réalité de référence, sont issues de mesures (cf. 2.2). Cette *réalité* n'est donc pas exempte d'erreurs et d'incertitudes, que nous avons estimées de l'ordre de 5 à 10 %. Et pourtant, c'est sans le moindre remords que nous ignorerons ici ces erreurs de mesure ! En effet, en l'état actuel des performances des prévisions météorologiques, les erreurs sur les observations restent bien inférieures à celles constatées sur les prévisions, au moins en ce qui concerne les prévisions de précipitations.

3.1.2.2 Choix d'une fonction de composition

Il est courant de considérer comme fonction h de composition des scores individuels une simple moyenne arithmétique :

$$h((s_i)_{i=1,N}) = \frac{1}{N} \sum_{i=1}^N s_i = \bar{s} \quad (3.2)$$

Toutefois, dans le cadre de l'utilisation opérationnelle d'un système de prévision, ce critère peut ne pas être le plus intéressant. En effet, lors de l'émission d'une prévision, on ne connaît pas *a priori* la valeur qui sera effectivement observée : on ne sait donc pas, au moment de l'utiliser, si la prévision est de bonne qualité. On ne sera donc enclin à lui faire confiance que s'il est rare que le système de prévision fasse de grosses erreurs. Il est ainsi préférable de favoriser les systèmes robustes en pénalisant volontairement les fortes erreurs.

¹La fonction S se définit généralement comme une distance entre les observations et les prévisions, donc plus cette distance est faible, meilleure est la prévision. Dans le cas où S est définie comme l'opposé ou l'inverse de cette distance (cela arrive parfois!), il faut bien sûr considérer que la meilleure prévision est associée aux plus grandes valeurs de S .

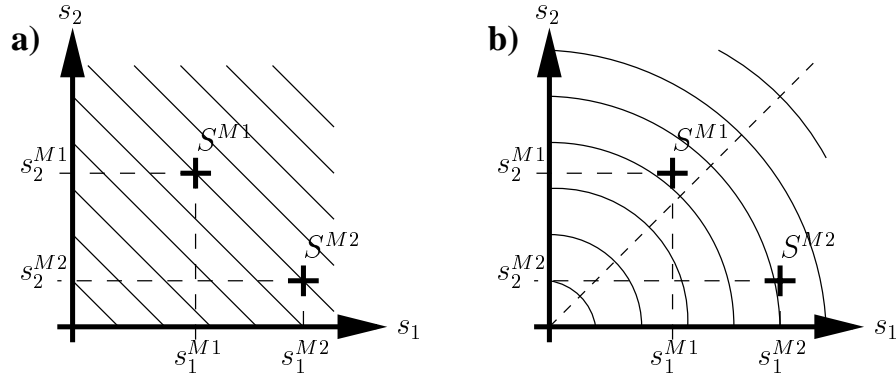


FIG. 3.1: Interprétation graphique de la fonction de composition h dans le cas d'une série de 2 prévisions individuelles pour deux algorithmes $M1$ et $M2$ ayant obtenu respectivement les séries de Scores Individuels $(s_1^{M1}; s_2^{M1})$ et $(s_1^{M2}; s_2^{M2})$. Sur un graphique à deux dimensions, le score S est représenté par un point de coordonnées $(s_1; s_2)$. **a)** Si h est une moyenne arithmétique : les lignes d'isovalues de S sont des droites. Dans ce cas, on a $S^{M1} = S^{M2}$. **b)** Si h est une moyenne quadratique : les lignes d'isovalues de S sont alors des arcs de cercle. On a ici, pour les mêmes données, un score global $S^{M1} < S^{M2}$ i.e. à moyenne arithmétique égale, l'algorithme $M1$ est considéré comme le plus performant, car c'est celui qui présente le plus petit écart-type de Score Individuel.

C'est pourquoi il peut être intéressant d'utiliser une moyenne quadratique en guise de fonction de composition h :

$$h\left((s_i)_{i=1,N}\right) = \sqrt{\frac{\sum_{i=1}^N s_i^2}{N}} = \sqrt{\bar{s}^2 + \sigma_s^2} \quad (3.3)$$

où σ_s^2 est la variance de la série (s_i) : $\sigma_s^2 = \frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})^2$.

Le Score Global S tient alors compte de la variance de la série en plus de la moyenne arithmétique. Il va donc favoriser, à performance moyenne égale, les prévisions dont les Scores Individuels présentent la plus faible dispersion. La qualité de chaque prévision individuelle reste dans ce cas toujours assez proche de la qualité moyenne, comme illustré par l'interprétation graphique présentée figure 3.1.

3.1.3 Attributs d'une prévision

La qualité d'une prévision repose sur sa capacité à posséder certaines propriétés désirables, appelées les *attributs* de la prévision (Epstein et Murphy 1965). Ceux-ci doivent donc être définis préalablement à la construction d'une mesure. En outre, même si les scores d'utilité sont construits indépendamment de ces attributs, il demeure intéressant de s'interroger sur leurs capacités à en rendre compte.

La définition de ces attributs fut l'objet de nombreuses publications, parfois contradictoires, dans les années 50 et 60. En souvenir de cette époque, certains attributs gardent plusieurs noms. Ainsi, la distorsion à petite échelle définie par Murphy et Epstein (1967), correspond-elle à la fiabilité définie par Sanders (1958), qui correspond elle-même à la justesse définie par Bross

(1953). Nous citerons ici les principaux attributs reconnus, qui semblent faire maintenant l'objet d'un consensus.

Ces attributs ont souvent été confondus avec les scores qui ont été utilisés pour les mesurer, mais à un même attribut peut correspondre plusieurs scores. Ceux-ci ne font que définir la métrique utilisée. C'est la même chose que pour mesurer par exemple une distance en plein cœur de Manhattan : si l'on dispose d'un hélicoptère, on mesurera la distance "à vol d'oiseau" (distance euclidienne), alors que si l'on se déplace par voie terrestre (en taxi pour faire couleur locale), il sera plus utile de mesurer la distance par la longueur à parcourir effectivement en empruntant les rues (distance de Manhattan ou city-block). Dans les deux cas on veut mesurer la même chose, la distance à parcourir, mais de façon différente.

3.1.3.1 La finesse (*sharpness* ou *resolution*)

La finesse correspond à la mesure de la similitude entre la prévision et une prévision catégorique (Bross 1953). Cet attribut mesure l'étendue du spectre des valeurs reconnues comme possibles par la prévision : cela correspond à la dispersion de la prévision.

Il convient de noter que la finesse d'une prévision est indépendante des observations. Seules les caractéristiques intrinsèques de la prévision influent sur la finesse.

Une prévision déterministe est par définition parfaitement fine : c'est une prévision catégorique. De même, une prévision probabiliste parfaitement fine est en fait une prévision catégorique. Dans le cas de variables bornées (telle l'humidité relative par exemple) il est possible de définir une borne supérieure au manque de finesse : c'est la prévision qui considère que l'on peut observer de façon équiprobable soit la plus petite soit la plus grande des valeurs admissibles (soit 0% et 100% pour l'humidité relative).

Cette notion de finesse s'illustre aisément graphiquement lorsque l'on considère des prévisions exprimées sous la forme de fonctions de répartition (Fig.3.2).

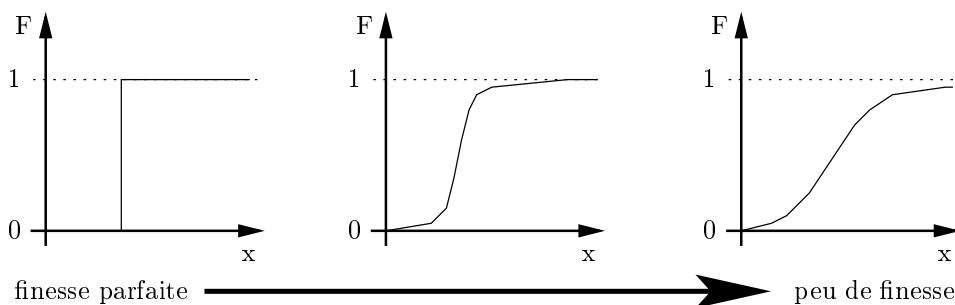


FIG. 3.2: Fonctions de répartition de trois prévisions différentes de la variable x . La première est de finesse parfaite (prévision déterministe), la seconde reste assez fine et la troisième montre peu de finesse.

3.1.3.2 La justesse (*accuracy* ou *validity*)

La justesse (Bross 1953; Epstein et Murphy 1965) est définie comme « le degré **moyen** de correspondance entre les prévisions individuelles et les observations » (Murphy et Daan 1985). En d'autres termes, elle représente la capacité d'une prévision à bien cerner la valeur qui sera observée *a posteriori* ; elle correspond au décalage entre prévisions et observations.

Dans le cas de la prévision catégorique d'une variable continue, la justesse est facile à définir. Elle se résume à la différence (ou la distance) entre la valeur observée et la valeur prévue.

Pour une prévision probabiliste, la justesse est plus délicate à cerner, car il n'y a plus une et une seule valeur prévue. Pour percevoir la justesse, il faut transformer la prévision probabiliste en prévision déterministe : soit par un critère arbitraire, en considérant par exemple que la prévision déterministe associée correspond à la médiane de la distribution, soit par un critère économique si l'utilisateur est identifié, en considérant alors une valeur optimale vis-à-vis d'une fonction de coût propre à cet utilisateur (cf. partie 3.3.5.1). On peut alors considérer la justesse comme la distance entre la valeur observée et la prévision déterministe issue de la prévision probabiliste. Cette définition peut paraître peu rigoureuse, mais c'est justement ce qui explique que cette notion de justesse ait donné lieu à tant de controverses.

La figure 3.3 illustre graphiquement cette notion en présentant, pour la même observation, deux prévisions de même finesse mais de justesse différente.

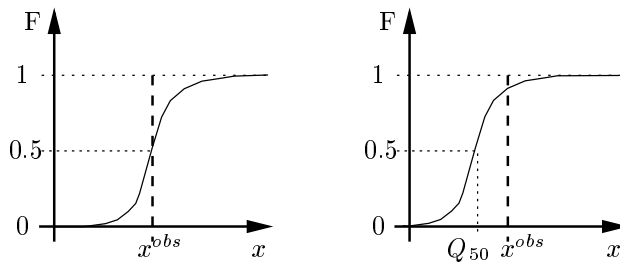


FIG. 3.3: Fonction de répartition de deux prévisions de même finesse, mais de justesse différente : la première est parfaitement juste alors que la seconde ne l'est pas (exemple dans le cas où l'on définit la justesse comme la distance entre la valeur médiane de la prévision Q_{50} et la valeur observée x^{obs}).

3.1.3.3 Le biais à grande échelle (*reliability* ou *bias*)

Lorsque l'on émet une prévision probabiliste, on fait l'hypothèse que l'observation correspondante sera une réalisation de la variable aléatoire associée à la loi de probabilité prévue. Si cette hypothèse est vérifiée, alors la prévision ne présente pas de biais. Par exemple, cela implique que les valeurs des quantiles 90% proposés pour chaque journée, n'ont réellement et exactement qu'une chance sur dix d'être dépassées lors de chaque couple prévision/observation. Cette propriété ne peut évidemment pas être vérifiée sur une seule réalisation. Nous allons donc parler de biais à grande échelle (Epstein et Murphy 1965), car il ne caractérise qu'un ensemble de réalisations suffisamment important (on croise aussi parfois le terme de *fiabilité*, mais celui-ci est

ambigu car selon les auteurs, il peut parfois, faire référence au biais à petite échelle dont nous parlons plus loin).

Plus généralement, la question que l'on se pose ici est de savoir si les quantiles $x\%$ n'ont effectivement pas été dépassés dans $x\%$ des cas par les observations sur l'ensemble des réalisations.

Considérons par exemple une prévision de probabilité d'occurrence de pluie réalisée sur un échantillon de N journées. Pour la facilité de l'exposé, la loi de probabilité prévue est supposée identique pour chacune des journées de l'échantillon ; celle-ci stipule que l'on a tous les jours une chance sur trois d'observer de la pluie :

- Si sur cet échantillon la probabilité empirique d'occurrence de pluie P_{emp} (pourcentage des journées lors desquelles on a observé des précipitations) est de $1/4$, alors la prévision est biaisée à grande échelle : ce qui a été prévu comme arrivant 1 fois sur 3 a en fait été observé 1 fois sur 4.
- Si $P_{emp} = 1/3$, alors la prévision n'est pas biaisée à grande échelle : elle est à même de reproduire la climatologie empirique.

Dans un cadre encore plus général, à partir d'une série de N prévisions \mathcal{P}_i qui cette fois ne sont plus toutes identiques, on calcule a posteriori, pour chaque élément i de la série de prévisions, la probabilité au non-dépassement P_i^{obs} associée à la valeur observée x_i^{obs} par la loi de probabilité prévue :

$$P_i^{obs} = \int_{-\infty}^{x_i^{obs}} f_i(x).dx \quad (3.4)$$

où f_i est la densité de probabilité donnée par la prévision \mathcal{P}_i (figure 3.4).

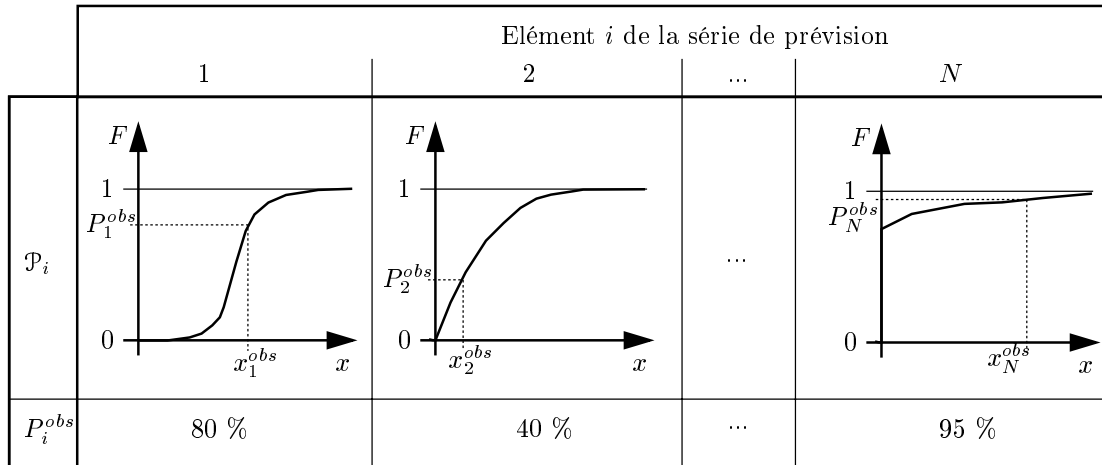


FIG. 3.4: Exemple de série de probabilités au non-dépassement P_i^{obs} obtenues à partir d'une série de prévisions \mathcal{P}_i .

On classe alors les P_i^{obs} , du plus petit au plus grand, en notant $P_{(i)}^{obs}$ l'élément classé n° i . On associe à chaque $P_{(i)}^{obs}$ sa fréquence empirique F_{emp} définie par $F_{emp} = \frac{(i)-0.375}{N+0.250}$ afin de tracer la fonction de répartition empirique des P_i^{obs} . Dans la littérature, un tel graphe est appelé *reliability chart*. La série de prévisions est non-biaisée si cette fonction suit la première bissectrice (Fig 3.5),

c'est-à-dire si la densité de probabilité des P_i^{obs} est uniforme. A l'inverse, si elle s'en écarte, le système de prévision est biaisé.

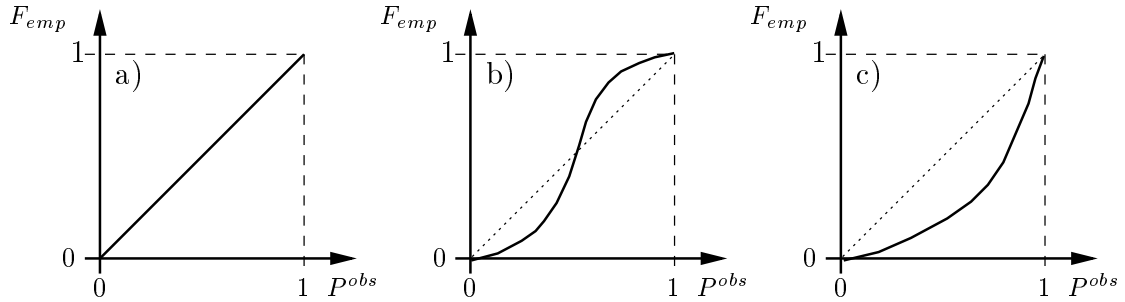


FIG. 3.5: Exemple de *reliability chart* (fonction de répartition de probabilité au non-dépassement prévue en vue de faire apparaître le biais à grande échelle). **a)** prévision non-biaisée : la fonction de répartition se confond avec la première bissectrice, **b)** prévision biaisée : les forts quantiles sont surestimés et les faibles sous-estimés, **c)** prévision biaisée : tous les quantiles sont sous-estimés.

Si l'on dispose d'un système de prévision présentant un biais à grande échelle identifié *a posteriori* sur un échantillon de prévisions, il est possible de le dé-biaser dans le but d'une utilisation ultérieure (Zhu *et al.* 1996). Pour cela, il suffit de remplacer chaque fonction de répartition prévue F_i par la fonction $(F_{emp} \circ F_i)$ où F_{emp} est, rappelons-le, la fonction de répartition empirique des P_i^{obs} sur l'échantillon. Cette correction nécessite toutefois de disposer d'un échantillon de prévisions suffisamment important avant sa mise en œuvre opérationnelle, afin d'être sûr que le biais soit un biais systématique de la méthode et non pas un biais dû à l'échantillonnage.

3.1.3.4 Le biais à petite échelle

Dans la partie précédente (3.1.3.3), nous avons postulé de façon péremptoire qu'il n'était pas possible de vérifier le biais d'une prévision isolée. En effet, nous ne disposons in fine que d'une réalisation de l'événement que l'on cherche à prévoir, la valeur observée, et malheureusement pas de sa loi de probabilité. Toutefois, si nous avons cette loi à notre disposition, nous pourrions parler de *biais à petite échelle* : l'incertitude contenue dans une prévision est-elle l'exacte traduction de l'incertitude du phénomène que l'on cherche à prévoir, toutes choses égales par ailleurs ?

Pour être capable de mesurer ce biais à petite échelle, il faudrait disposer d'un nombre important de prévisions toutes identiques entre elles et des observations (peut-être toutes différentes) qui leur correspondent. De cette façon, nous aurions accès à la loi de probabilité empirique des observations pour cette prévision (loi conditionnelle au fait qu'elles ont toutes la même prévision probabiliste a priori). On pourrait alors la comparer à celle donnée par la prévision.

Malheureusement, recueillir plusieurs prévisions exactement identiques est hautement irréaliste, et le biais à petite échelle demeure inaccessible. Cependant, il reste possible de déterminer si un score de vérification favorise ou non les prévisions biaisées à petite échelle. Ce faisant, nous quittons le domaine des attributs des prévisions à proprement parler pour entrer dans celui des attributs des scores de vérification.

Nous définissons donc les **scores non-biaisés**. Considérons ainsi un ensemble de N observations x_i^0 qui définissent une loi de probabilité empirique \mathcal{L}_{Clim} , climatologie propre de cet échantillon. Un score sera dit non-biaisé si, appliqué à une loi de probabilité prévue \mathcal{L}_{prevue} , unique ici sur tout l'échantillon, il se révèle optimal¹ pour $\mathcal{L}_{prevue} = \mathcal{L}_{Clim}$.

3.1.4 Tout est relatif ...

La mesure des différents attributs d'une prévision n'est pas suffisante pour dire qu'une série de prévisions est bonne ou mauvaise. Il est nécessaire pour cela de comparer ces mesures à une (ou des) valeur(s) de référence.

Ainsi, il est aisé de réaliser des prévisions de précipitation parfaitement fines et exactes dans 66% des cas sur le groupement pluviométrique de Corse du sud. Exprimé de cette façon, ces prévisions peuvent sembler très intéressantes et relever d'une belle prouesse scientifique ! Or, pour arriver à ces performances, il suffit simplement de prévoir tous les jours qu'il ne pleuvra pas. En effet, il ne pleut en moyenne qu'un jour sur trois sur ce groupement. Ces performances ne sont donc pas meilleures que ce que l'on peut obtenir simplement en connaissant la climatologie du bassin : elles ne font qu'atteindre le minimum que l'on est en droit d'attendre d'une prévision.

Les mesures de qualité des prévisions seront donc toujours comparées aux mesures de qualité d'une prévision de référence. On est ainsi amené à définir la Compétence (*Skill Score* dans la littérature en langue anglaise) $\mathbf{Comp}_S(M_1)$ de la méthode de prévision M_1 , basée sur le Score Global S :

$$\mathbf{Comp}_S(M_1) = \frac{S(M_1) - S(M_{Ref})}{S_{parfait} - S(M_{Ref})} \quad (3.5)$$

où M_{Ref} est la méthode de prévision de référence et $S_{parfait}$ est la valeur du score S obtenue par une prévision parfaite. Une compétence de 1 correspond à une prévision parfaite, de 0 à une prévision équivalente à la prévision de référence, et une compétence négative correspond à une prévision de moins bonne qualité que la prévision de référence.

La référence naturelle en météorologie est l'espérance climatologique, et c'est celle que nous allons utiliser tout au long de ce mémoire. Ainsi, pour les cumuls de précipitations, notre prévision de référence sur un groupement correspondra à la fonction de répartition de toutes les pluies recueillies sur ce groupement. Bien que cela eût été plus rigoureux, nous avons choisi de ne pas différencier les climatologies selon les saisons et de simplement nous limiter à une climatologie annuelle globale par groupement, évitant ainsi d'augmenter trop les temps de calcul. Pour un groupement donné, la référence sera donc la même quelle que soit la date de la prévision.

Notons qu'il arrive souvent en météorologie que l'on prenne la *persistance* comme prévision de référence : la prévision de référence est alors l'observation correspondant au pas de temps précédent. Les variables sensibles qui intéressent la prévision étant généralement autocorrélées, la persistance est une méthode basique de prévision plus efficace que la climatologie. Elle n'est prise en défaut qu'à chaque changement de temps, mais c'est justement ce que l'on cherche à prévoir ! Nous ne l'avons cependant pas retenue ici car la persistance fournit une prévision

¹c'est à dire minimum s'il est orienté négativement, et maximum sinon.

déterministe alors que nous allons principalement travailler avec des prévisions probabilistes ; l'usage de la climatologie comme méthode de référence permet de rester homogène en comparant deux prévisions probabilistes. De toute manière, le choix d'une méthode de référence ou d'une autre est sans incidence sur le classement relatif de deux systèmes de prévision, ce qui constituera notre principal objectif.

Maintenant que nous avons défini notre compétence, nous avons une obligation de résultat. Il nous faut prouver l'habileté de nos prévisions, car comme l'a écrit Sanders (1963) : si les prévisions se montrent incompetentes, l'utilisateur des avis météorologiques est en droit de se demander si les moyens mis en œuvre sont justifiables et s'ils méritent d'être maintenus !

3.2 Qualité des prévisions déterministes

Comme nous l'avons vu précédemment, une prévision déterministe est une prévision qui n'affiche pas d'incertitudes et fournit une valeur unique à l'utilisateur (cf. 1.1.3). Il n'y a donc pas lieu de mesurer de finesse pour ces prévisions, celles-ci étant toujours de finesse parfaite. Les scores de qualité utilisés sont donc des scores qui mesurent soit une justesse, soit un biais à grande échelle.

3.2.1 Prévision par catégories

3.2.1.1 Prévision oui/non

Tableaux de contingence

Lorsque la prévision est utilisée à des fins d'alerte, on est généralement intéressé par le dépassement ou non d'un certain seuil par une variable sensible, comme par exemple, un niveau d'eau en amont d'une agglomération dans le cas d'une alerte de crue. Tant que cette variable reste en deçà de ce seuil, tout va pour le mieux. Cependant, si elle le dépasse, on peut craindre des dommages aux biens ou aux personnes et il est nécessaire de déclencher une alerte.

De même, la prévision peut concerner l'occurrence ou non d'un phénomène météorologique, sans s'intéresser pour autant à son amplitude : occurrence de cyclone ou de précipitation par exemple.

Il est donc courant de ranger observations et prévisions au sein de deux catégories :

- d'une part les *événements réalisés* (dépassement d'un certain seuil d'alerte X_s prédéfini, occurrence du phénomène météorologique suivi, etc ...),
- et d'autre part les *événements non réalisés*.

Dans toute la suite, par souci de clarté, nous considérerons qu'un événement réalisé est une alerte.

A partir de ce classement binaire, on peut construire un tableau de contingence (tableau 3.2) en comptant les alertes observées/non observées (Oui/Non) en fonction des alertes prévues/non prévues (Oui/Non).

		Observations	
		Oui	Non
Prévisions	Oui	a	b
	Non	c	d

TAB. 3.2: Tableau de contingence dans le cas d'une prévision en deux catégories. Oui correspond à l'occurrence de l'alerte (dépassement de seuil, présence d'un cyclone, pluie non nulle...) et Non à l'absence d'occurrence de l'alerte.

Ce tableau permet de définir ainsi :

- a : nombre d'alertes observées effectivement prévues (bonne alerte) ;
- b : nombre d'alertes prévues mais non observées (fausse alerte) ;
- c : nombre d'alertes observées mais non prévues (défaut d'alerte) ;
- d : nombre d'observations n'ayant donné lieu ni à une alerte ni à une prévision d'alerte ('NON' correct) ;
- n : nombre total d'observations sur lesquelles porte la prévision : $n = a + b + c + d$.

On remarque que les quantités $a + c$ et $b + d$ (sommées selon les colonnes) sont imposées par la climatologie de l'événement sur l'échantillon. Ainsi, $(a + c)/n$ correspond à la fréquence empirique d'alertes observées. A l'inverse, les sommes selon les lignes sont imposées par la méthode de prévision : $(a + b)/n$ correspond à la fréquence des alertes prévues.

Scores de vérification

Pour une bonne méthode de prévision, b et c doivent être les plus faibles possibles. Cependant, selon que l'on accorde plus ou moins d'importance aux fausses alertes ou aux défauts d'alerte, selon que l'on prend en compte la climatologie des événements, ou autre, il est possible de considérer différents scores de qualité. La multiplicité de ces scores (associée à la multiplicité des appellations recouvrant un même score) entraîne une certaine confusion dans la littérature. Ainsi, FAR et $1 - SPE$ (cf. ci-dessous) sont souvent confondus entre eux comme le fait remarquer Golding (2000). Nous essayons de présenter ici les scores usuels de façon cohérente.

- **Indice de Réussite (*Hit Rate*, H) :**

$$H = \frac{a + d}{n}$$

L'Indice de Réussite correspond au nombre de prévisions qui se sont révélées justes, rapporté au nombre de prévisions émises. Ce score ne fait pas de distinction entre fausse alerte et défaut d'alerte, et ne prend pas en compte la climatologie de l'événement sur l'échantillon.

- **Indice de Succès Critique (*Threat Score*, TS ou *Critical success index*, CSI) :**

$$CSI = \frac{a}{a + b + c}$$

L'Indice de Succès Critique correspond au nombre de bonnes alertes rapporté au nombre d'événements pour lesquels une alerte a été prévue ou aurait dû être prévue. Ce score

est intéressant lorsque la fréquence empirique des événements donnant lieu à des alertes est faible devant celle des événements sub-seuil. Il est en effet facile dans ce cas de faire fréquemment une bonne prévision en ne prévoyant jamais d'alerte. Or, ce qui est attendu de la prévision, c'est d'être efficace dans le cas des alertes, que celles-ci soient prévues ou observées. Le score *CSI* permet d'ignorer les 'NON' corrects.

- **Score de Menace Equitable** (*Equitable Threat Score, ETS*) :

$$ETS = \frac{a - R}{a + b + c - R} \quad \text{où} \quad R = \frac{(a + c) \cdot (a + b)}{n}$$

Le terme R est l'espérance du nombre de bonnes alertes faites en émettant de manière totalement aléatoire $(a + b)$ prévisions d'un événement (le 'oui') de fréquence $(a + c)/n$. Il correspond donc au nombre de bonnes alertes réalisées « *par chance* ». Ce score est une modification du *CSI* : en omettant les bonnes prévisions dues au hasard, il permet de comparer entre elles des prévisions émises pour des régimes climatiques différents.

- **Probabilité de Détection** (*Probability Of Detection, POD*) :

$$POD = \frac{a}{a + c}$$

La Probabilité de Détection est égale au nombre de bonnes alertes sur le nombre total d'alertes observées. Cela correspond à la probabilité, lorsqu'une alerte a effectivement lieu, que celle-ci ait été préalablement prévue. C'est donc la probabilité conditionnelle de faire une bonne prévision d'alerte, sachant qu'une alerte est observée. Ce score ignore totalement les possibilités de fausses alertes.

- **Taux de Fausses Alertes** (*False Alarm Rate, FAR*) :

$$FAR = \frac{b}{a + b}$$

Ce score est le pendant du précédent. Il est égal au nombre de fausses alertes émises rapporté au nombre total d'alertes prévues. Cela correspond à la probabilité, lorsqu'une alerte est prévue, que celle-ci n'ait pas lieu d'être. C'est donc ici aussi une probabilité conditionnelle : c'est la probabilité de faire une mauvaise prévision sachant qu'une alerte a été prévue. Ce score ignore pour sa part totalement les possibilités de défaut d'alerte. Contrairement aux scores présentés précédemment, le *FAR* est orienté négativement, c'est-à-dire qu'une prévision parfaite présente un *FAR* de 0% alors que le plus mauvais score possible est de 100%.

- **Spécificité** (*Specificity, SPE*) :

$$SPE = \frac{d}{b + d}$$

Surtout employée en médecine, la Spécificité correspond à la probabilité conditionnelle d'émettre un 'NON' correct lorsque qu'aucune alerte n'est effectivement observée. On définit par ailleurs le Taux de Fausse Détection comme étant égal à $1 - SPE$.

– **Biais** (*Bias*, B) :

$$B = \frac{a + b}{a + c}$$

Le Biais correspond au rapport entre les fréquences prévues et empiriques d'alerte. Il ne préjuge pas de la qualité de la prévision, mais plutôt de sa cohérence climatologique. En effet, rien n'indique dans ce score que les alertes prévues et observées soient concomitantes.

Afin d'illustrer l'utilisation de ces différents scores, considérons la mise en œuvre de trois méthodes de prévision différentes, M_1 , M_2 et M_3 , sur un même échantillon d'événements. Les tableaux de contingences calculés à partir des prévisions issues de ces trois méthodes sont donnés par le tableau 3.3.

a)		Observations		b)		Observations		c)		Observations	
		Oui	Non			Oui	Non			Oui	Non
Prévisions	Oui	0	0	Prévisions	Oui	6	6	Prévisions	Oui	10	90
	Non	10	90		Non	4	84		Non	0	0

TAB. 3.3: Exemple de tableaux de contingences obtenus à partir de méthodes de prévisions différentes sur le même échantillon d'événements. **a)** méthode M_1 , **b)** méthode M_2 , **c)** méthode M_3 .

	$H(\%)$	$CSI(\%)$	$ETS(\%)$	$POD(\%)$	$FAR(\%)$	$SPE(\%)$	$Bias$
M_1	90	0	0	0	0	100	0.0
M_2	90	38	32	60	50	93	0.8
M_3	10	10	0	100	90	0	10.0

TAB. 3.4: Calcul de différents scores pour les méthodes de prévisions M_1 , M_2 et M_3

La méthode M_1 ne prévoit jamais la moindre alerte alors que M_3 prévoit une alerte tous les jours. La méthode M_2 , quant à elle, ose s'engager plus en essayant de faire réellement une prévision d'alerte. Le calcul des différents scores est indiqué par le tableau 3.4. Il apparaît que les méthodes M_1 et M_2 présentent le même Indice de Réussite $H = 90\%$. Ce score est essentiellement dû aux 'NON' corrects et cet artifice donne l'impression que la méthode M_1 est très performante. Les scores CSI et ETS permettent d'établir une meilleure hiérarchie. L'Indice de Réussite H est donc un score peu utile : il est plus trompeur qu'informatif. Il est d'ailleurs étonnant que nombre de services météorologiques persistent à quantifier leurs performances de prévisions d'occurrence de pluie de cette façon, atteignant de cette manière des taux de réussite allant jusqu'à 89%. Cela impressionne certes le grand public, mais n'est pas réellement probant d'un point de vue scientifique.

On peut aussi remarquer que de bons scores en Probabilité de Détection ou en Taux de Fausse Alerte ne sont pas forcément synonymes de bons scores en Indice de Succès Critique. Ces deux

attributs, *POD* et *FAR* sont souvent antagonistes et l'utilisateur peut être amené à privilégier l'un ou l'autre en fonction des objectifs liés à la mise en œuvre de la prévision.

Scores de compétence

En se fixant une méthode de prévision de référence M_{Ref} , il est bien sûr possible de calculer des scores de compétence à partir de ces scores de vérification (cf. 3.1.4). On rappelle que la compétence $\mathbf{Comp}_S(M_1)$ d'une méthode de prévision M_1 , basée sur le Score Global S , est définie par :

$$\mathbf{Comp}_S(M_1) = \frac{S(M_1) - S(M_{Ref})}{S_{parfait} - S(M_{Ref})} \quad (3.6)$$

Certains de ces scores de compétence sont utilisés de façon classique et ont reçu des dénominations propres dans la littérature :

- **Score de Heidke (*Heidke skill score, HSS*)** : Le score *HSS* correspond à la compétence de l'Indice de Réussite H par rapport à une méthode de prévision M_{Ref} fournissant une prévision \mathcal{P}_{hasard} faite au hasard mais selon la même fréquence de prévisions d'alertes. Il est donc défini par :

$$HSS = \frac{H - H_{hasard}}{1 - H_{hasard}} \quad \text{avec} \quad H_{hasard} = \frac{a_{hasard} + d_{hasard}}{n}$$

$$\text{où : } \left\{ \begin{array}{l} H_{hasard} \text{ est l'Indice de Réussite des prévisions issues de } \mathcal{P}_{hasard}, \\ a_{hasard} = \frac{(a+b).(a+c)}{n} \text{ est l'espérance du nombre de bonnes prévisions émises par } \mathcal{P}_{hasard}, \\ d_{hasard} = \frac{(b+d).(c+d)}{n} \text{ est l'espérance du nombre de 'NON' corrects émis par } \mathcal{P}_{hasard}. \end{array} \right.$$

Le développement analytique du score *HSS* conduit à :

$$HSS = \frac{2.(ad - bc)}{(a + c).(c + d) + (a + b).(b + d)}$$

- **Score d'Hanssen-Kuipers (*True skill score, TSS*)** : Le score *TSS* n'est pas à proprement parler un score de compétence, mais il en est très proche. Il correspond à la différence d'Indice de Réussite entre la prévision testée et \mathcal{P}_{hasard} mais rapportée à l'Indice de Réussite obtenu par une prévision \mathcal{P}_{clim} faite elle aussi au hasard mais respectant la climatologie observée de l'événement sur l'échantillon considéré :

$$TSS = \frac{H - H_{hasard}}{1 - H_{clim}} \quad \text{avec} \quad H_{clim} = \frac{a_{clim} + d_{clim}}{n}$$

$$\text{où : } \left\{ \begin{array}{l} H_{clim} \text{ est l'Indice de Réussite des prévisions issues de } \mathcal{P}_{clim}, \\ a_{clim} = \frac{(a+c)^2}{n} \text{ est l'espérance du nombre de bonnes prévisions émises par } \mathcal{P}_{clim}, \\ d_{clim} = \frac{(b+d)^2}{n} \text{ est l'espérance du nombre de 'NON' corrects émis par } \mathcal{P}_{clim}. \end{array} \right.$$

Le score TSS s'exprime donc finalement par :

$$TSS = \frac{ad - bc}{(a + c).(b + d)}$$

Lorsque l'on est en présence d'une prévision sans biais à grande échelle ($Bias = 1$), le score d'Hanssen-Kuipers est équivalent au score de Heidke. Par ailleurs, le score d'Hanssen-Kuipers peut s'exprimer à partir des scores POD et SPE définis précédemment :

$$TSS = POD - (1 - SPE)$$

Du fait de la confusion qui règne généralement entre FAR et $1 - SPE$, il arrive que le score TSS soit maladroitement calculé comme étant égal à $POD - FAR$. Il convient donc de rester vigilant lors de son utilisation.

- **Indice de Rousseau (I)** : L'Indice de Rousseau a été présenté initialement à l'aide d'une interprétation graphique des tableaux de contingences (Rousseau 1980) :

$$I = \frac{ad - E^2}{(a + E).(d + E)} \quad \text{avec} \quad E = \frac{b + c}{2}$$

On peut toutefois l'expliquer comme un score de compétence. Pour ce faire, on considère, non plus la prévision réellement obtenue, mais la prévision \mathcal{P}_R de même Indice de Réussite, même Indice de Succès Critique et telle que $Bias = 1$, *i.e.* telle que le biais à grande échelle soit nul. Le tableau de contingence de la prévision \mathcal{P}_R est alors tel que présenté ci-dessous. L'Indice de Rousseau correspond alors au Score de Heidke obtenu par cette prévision \mathcal{P}_R .

		Observations	
		Oui	Non
Prévisions	Oui	a	E
	Non	E	d

Le tableau 3.5 illustre l'utilisation de ces scores de compétence sur les données des exemples présentés par le tableau 3.3. Les trois scores de compétence montrent que le fort Indice de Réussite de la méthode M_1 est tout à fait artificiel. Seul l'Indice de Rousseau arrive à établir une hiérarchie sur les trois méthodes à la fois : la méthode M_3 est considérée comme encore moins performante que M_1 du fait de son fort biais.

	$H(\%)$	$HSS(\%)$	$TSS(\%)$	$I(\%)$
M_1	90	0	0	-5
M_2	90	49	53	49
M_3	10	0	0	-82

TAB. 3.5: Calcul de différents scores de compétence pour les méthodes de prévisions M_1 , M_2 et M_3 .

3.2.1.2 Prévisions multicatégories

Il peut arriver que l'on ait besoin de plus de deux catégories pour “ranger” les événements auxquels on s'intéresse :

- Si l'on surveille par exemple le niveau d'une rivière, on peut définir deux valeurs seuils : une première valeur X_{s1} qui déclenche une mise en vigilance, et une seconde X_{s2} qui déclenche une alerte. On rangera alors les événements en trois catégories : ceux qui sont inférieurs aux deux seuils (catégorie A : aucune action à entreprendre), ceux qui sont compris entre ces deux seuils (catégorie B : mise en vigilance) et enfin ceux qui sont supérieurs aux deux seuils (catégorie C : déclenchement de l'alerte). Dans ce cas, les différentes catégories sont ordonnées entre elles.
- Si l'on s'intéresse à la prévision d'occurrence de précipitation, il peut être intéressant de faire un distinguo entre précipitation liquide et précipitation solide. On peut ainsi définir quatre catégories par exemple : aucune précipitation (A), occurrence de pluie (B), occurrence de neige (C) et occurrence de grêle (D). Les différentes catégories ne sont pas ordonnées entre elles dans ce cas.

Nous insistons sur le fait que l'on ne considère ici que des prévisions déterministes, et que la prévision consiste à annoncer une et une seule catégorie. On calcule alors un tableau de contingence analogue à celui présenté dans le tableau 3.2, mais présentant plus d'entrées. Le tableau 3.6 propose un tel tableau dans un cas à trois catégories.

		Observations		
		A	B	C
Prévisions	A	a	b	c
	B	d	e	f
	C	g	h	i

TAB. 3.6: Tableau de contingence dans le cas d'une prévision en trois catégories.

Seul l'Indice de Réussite est facilement généralisable au contexte multicatégorie. Il est alors égal à la somme des éléments diagonaux du tableau (du coin en haut à gauche au coin en bas à droite) divisée par la somme de tous les éléments du tableau. Cependant, comme on l'a vu précédemment, même si ce score est le plus intuitif, il est loin d'être le plus pertinent. Pour calculer les autres scores (CSI , POD , FAR ...) il est plus aisé de se ramener au cas à deux catégories. Pour chacune des catégories, on calcule un tableau de contingence Oui/Non en considérant que Oui correspond à la catégorie cible, et Non à l'ensemble de toutes les autres catégories. On calcule ainsi autant de tableaux de contingence Oui/Non que l'on a initialement de catégories dans le tableau de contingence multicatégorie (tableau 3.7). Les différents scores peuvent alors être calculés sur chacun des tableaux de contingence Oui/Non. Si les catégories d'événements sont ordonnées entre elles, il est possible de tracer l'évolution de ces scores en

		Observations	
		A	Non A
Prévisions	A	a	$b + c$
	Non A	$d + g$	$e + f + h + i$

		Observations	
		B	Non B
Prévisions	B	e	$d + f$
	Non B	$b + h$	$a + c + g + i$

		Observations	
		C	Non C
Prévisions	C	i	$g + h$
	Non C	$c + f$	$a + b + d + e$

TAB. 3.7: Exemple de ré-écriture du tableau de contingence 3.6 en trois tableaux de contingence Oui/Non.

fonction des valeurs des seuils qui délimitent les catégories. De tels graphes donnent une vision à la fois synthétique et très détaillée de la performance de la prévision.

3.2.2 Prévision continue

Nombre de variables météorologiques sont des variables quantitatives continues : cumul de précipitation en un temps donné, température, humidité, etc. Contrairement au cas précédent, dans lequel les événements ne pouvaient prendre qu'un nombre fini de valeurs (les catégories), on se trouve ici confronté à des variables susceptibles de prendre un nombre infini de valeurs, toutes ordonnées entre elles. Ainsi, la prévision « *demain il pleuvra* » est une prévision en catégorie (pluie/non-pluie) alors que la prévision « *demain il pleuvra 10 mm* » est une prévision continue.

3.2.2.1 Agrégation en catégories

Une première approche pour calculer la qualité de ce type de prévision consiste à se ramener au cas précédent en transformant la prévision continue en prévision en catégorie.

Pour ce faire, on définit une valeur seuil X_s et on sépare les événements en deux groupes : d'un côté ceux qui dépassent le seuil, et de l'autre côté ceux qui lui sont inférieurs. On peut ainsi construire un tableau de contingence analogue à celui présenté par le tableau 3.2 sur lequel il est possible de calculer les scores qui nous intéressent. En faisant varier le seuil X_s et en recommençant l'opération, on peut ainsi obtenir l'évolution des scores de vérification choisis en fonction de la valeur du seuil.

3.2.2.2 Scores d'erreur moyenne

Erreurs calculées sur les valeurs

Lorsque la prévision est continue, les différentes valeurs que peuvent prendre les observations et les prévisions sont ordonnées entre elles. Il est donc possible de calculer une distance séparant chaque prévision de chaque observation, définissant ainsi un Score Individuel.

Il est courant de considérer un Score Individuel s_i égal à l'écart absolu entre observation et prévision :

$$s_i = |x_i^* - x_i^0| \quad (3.7)$$

où : $\left| \begin{array}{l} x_i^* \text{ est la } i^{\text{ème}} \text{ prévision de la série,} \\ x_i^0 \text{ est la } i^{\text{ème}} \text{ observation de la série.} \end{array} \right.$

A partir de ce Score Individuel, on définit les Scores Globaux $RMSE$ (Root Mean Squared Error) et MAE (Mean Absolute Error), selon que l'on choisit comme fonction de composition h une moyenne respectivement quadratique ou arithmétique :

$$RMSE = \sqrt{\frac{\sum_{i=1}^N s_i^2}{N}} = \sqrt{\frac{\sum_{i=1}^N (x_i^* - x_i^0)^2}{N}} \quad (3.8)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N s_i = \frac{1}{N} \sum_{i=1}^N |x_i^* - x_i^0| \quad (3.9)$$

Ces scores sont généralement utilisés pour quantifier la qualité des prévisions de variables telles que les champs d'altitude de géopotentiels. Ils permettent de mesurer un écart moyen entre les prévisions et les observations mais n'indiquent pas clairement si ces deux séries coïncident.

Erreurs calculées sur les probabilités empiriques

Si la variable sur laquelle porte la prévision déterministe est un cumul de précipitations, il semble logique de considérer que faire 10 mm d'erreur sur un événement de 100 mm n'a pas la même signification que sur un événement de 20 mm. En effet, les fortes valeurs de précipitations sont moins représentées que les faibles valeurs. Ainsi, dans la distribution climatologique, l'écart en fréquence empirique entre 110 et 100 mm est largement plus petit qu'entre 30 et 20 mm.

Ainsi, pour des variables qui ne sont pas réparties de façon homogène dans l'ensemble des valeurs admissibles, il peut sembler intéressant de considérer non pas l'erreur sur les valeurs prises par les variables (prévue vs. observée), comme dans le cas du $RMSE$ et du MAE , mais plutôt celle commise sur les fréquences empiriques cumulées associées à ces valeurs. On définit de cette manière un nouveau Score Individuel s_i (fig. 3.6) :

$$s_i = |F(x_i^*) - F(x_i^0)| \quad (3.10)$$

où :

- F est la fonction de répartition empirique de la variable x , établie sur une archive climatologique,
- x_i^* est la $i^{\text{ème}}$ prévision de la série,
- x_i^0 est la $i^{\text{ème}}$ observation de la série.

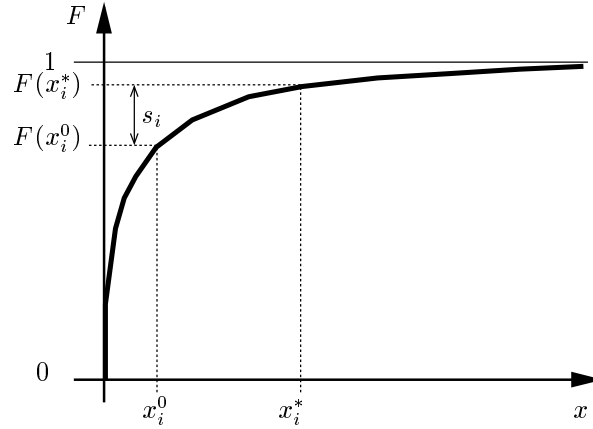


FIG. 3.6: Exemple de fonction de répartition cumulée empirique F d'une variable x .

A partir de ce Score Individuel, on établit le Score Global *LEPS* (Linear Error in Probability Space : Ward et Folland 1991 ; revisité par Potts *et al.* 1996, qui en donnent une version plus évoluée) :

$$LEPS = \frac{1}{N} \sum_{i=1}^N s_i = \frac{1}{N} \sum_{i=1}^N |F(x_i^*) - F(x_i^0)| \quad (3.11)$$

Ce score est principalement utilisé pour juger des prévisions saisonnières, pour lesquelles ce qui importe est moins l'intensité des valeurs prévues que leur place dans la climatologie moyenne.

Ce score est typiquement un score de qualité. En effet, il pénalise moins les erreurs sur les valeurs rares que sur les valeurs centrales. Or, ce n'est pas parce qu'une valeur est rare qu'une erreur sur son estimation est forcément anodine pour l'utilisateur de la prévision. L'utilité d'une prévision n'a donc aucune raison de correspondre au score *LEPS*.

3.2.2.3 Analyse en corrélation

Une analyse en corrélation permet de quantifier la co-fluctuation entre les valeurs prévues et observées. Autrement dit, quand ces dernières possèdent de fortes ou de faibles valeurs, en est-il de même pour les prévisions ?

L'explication complète de la théorie de la régression linéaire sortirait très largement du cadre de ce mémoire. Le lecteur désirant approfondir les notions présentées ici est donc invité à consulter le très pédagogique cours des professeurs Bois et Obled (2001) ou encore l'ouvrage de Draper et Smith (1981) s'il est anglophone.

Le score Global résultant d'une analyse en corrélation est le coefficient de corrélation r . On l'utilise souvent élevé au carré : r^2 , et l'on parle alors de coefficient de détermination.

Pour calculer ces scores, sur une série de N prévisions, on considère le nuage de points obtenu en portant sur un graphique les N points de coordonnées (x_i^*, x_i^0) (en reprenant les notations définies précédemment). Si l'on trace une droite dans ce nuage, il est possible de calculer la distance verticale e_i qui sépare chaque point de la droite (figure 3.7). On cherche alors quelle est la droite qui minimise la moyenne quadratique des e_i : c'est la droite d'estimation des observations à partir des prévisions. Le coefficient de corrélation est alors lié au rapport entre la variance s_e^2 des e_i définis par rapport à cette droite et la variance $s_{x^0}^2$ de la série des observations x_i^0 :

$$r^2 = 1 - \frac{s_e^2}{s_{x^0}^2} \quad (3.12)$$

Le coefficient de détermination correspond donc à la part de la variance des observations qui est « expliquée » par les prévisions.

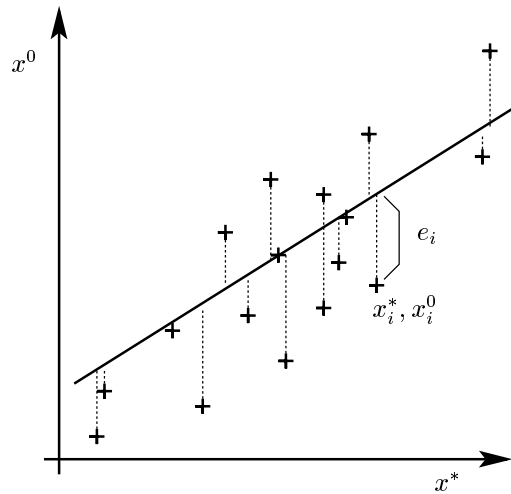


FIG. 3.7: Schéma d'un nuage de points au milieu duquel passe une droite : calcul des distances e_i (d'après Bois et Obled, 1998)

Dans la pratique, r se calcule grâce à l'équation suivante :

$$r = \frac{\frac{1}{N-1} \cdot \sum_{i=1}^N (x_i^0 - m_{x^0})(x_i^* - m_{x^*})}{s_{x^0} \cdot s_{x^*}} \quad (3.13)$$

où :

$$\left| \begin{array}{l} m_{x^0} \text{ est la moyenne arithmétique des observations : } m_{x^0} = \frac{1}{N} \cdot \sum_{i=1}^N x_i^0 \\ m_{x^*} \text{ est la moyenne arithmétique des prévisions : } m_{x^*} = \frac{1}{N} \cdot \sum_{i=1}^N x_i^* \\ s_{x^0} \text{ est l'écart-type des observations : } s_{x^0} = \sqrt{\frac{1}{N-1} \cdot \sum_{i=1}^N (x_i^0 - m_{x^0})^2} \\ s_{x^*} \text{ est l'écart-type des prévisions : } s_{x^*} = \sqrt{\frac{1}{N-1} \cdot \sum_{i=1}^N (x_i^* - m_{x^*})^2} \end{array} \right.$$

Bien que l'équation 3.13 permette de calculer directement le score r , il est toujours intéressant de tracer le nuage de points sur lequel est effectuée l'analyse en corrélation. Ceci permet en effet d'éviter certains pièges :

- Quelques points très éloignés du nuage principal peuvent conduire à une surestimation non négligeable du coefficient de corrélation (figure 3.8 a). Il en est de même si l'on se trouve en présence d'un point d'accumulation (pluies nulles par exemple dans le cas de prévisions de précipitations).
- Les coefficients de corrélation et de détermination ne donnent pas directement de renseignements quant au biais systématique possible entre prévision et observation. Ainsi, deux séries peuvent présenter un score r très important, même si l'écart entre ces séries est systématiquement d'un facteur 3 (figure 3.8 b). Il est donc recommandé d'utiliser l'analyse en corrélation en relation avec un score qui mesure l'écart moyen entre les séries, tel que $RMSE$ ou MAE .

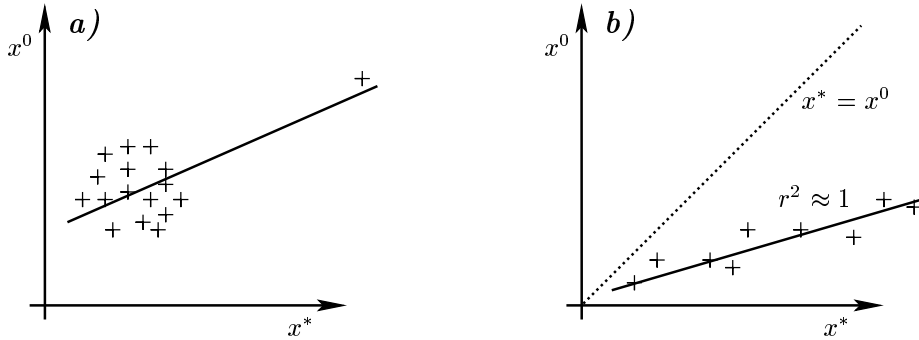


FIG. 3.8: Exemples de nuages de points servant à caler une droite de régression.

3.3 Utilité des prévisions probabilistes

Jusqu'ici, nous considérons des prévisions déterministes, caractérisées par une seule valeur prévue pour chaque événement. Or, comme nous l'avons vu dans la partie 1.1.3, les prévisions peuvent aussi être émises sous forme probabiliste. Celles-ci se présentent généralement sous la forme d'une fonction de répartition conditionnelle si elles portent sur une variable continue, ou sous la forme d'un ensemble de probabilités d'appartenance à différentes catégories si la variable est discrète.

3.3.1 Spécificité des prévisions probabilistes

La principale difficulté à laquelle on se trouve confronté, lorsque l'on veut définir la qualité d'une prévision probabiliste, réside dans le fait que l'on doit comparer deux objets de nature différente (Wilson *et al.* 1996, cité par Talagrand *et al.* 1997). En effet, la prévision est de nature probabiliste alors que la réalisation est pour sa part déterministe.

Il n'est donc pas possible de qualifier une telle prévision, prise isolément, de *bonne* ou *mauvaise*. Une prévision probabiliste associe en général une probabilité non nulle à n'importe quelle réalisation : cela revient à dire que tout est possible ! Ainsi, aussi faible que soit la probabilité associée à la valeur de la réalisation effective, il faut bien admettre que celle-ci a été prévue.

Il n'y a donc pas de sens à parler de la qualité d'une prévision probabiliste isolée. C'est pourquoi il nous est interdit, dans ce cas, de considérer un Score Individuel s mesurant la qualité. Certains auteurs, tels Talagrand *et al.* (1997), considèrent alors que l'on doit se contenter de s'intéresser uniquement à la qualité de séries de prévisions probabilistes, en utilisant un Score Global S qui n'est plus fondé sur la composition de Scores Individuels. On compare alors, sur un échantillon suffisamment long, une série de prévisions avec une série d'observations, ce qui nous ramène dans une situation plus confortable intellectuellement parlant.

Toutefois, même s'il n'est pas possible de parler de la qualité d'une prévision probabiliste prise isolément, il est possible (et même conseillé) de parler de son utilité. Dans tous les cas, la prévision est peut-être juste, mais est-elle utile pour autant ?

L'utilité n'est pas basée sur la différence entre la prévision et l'observation, contrairement à la qualité, mais sur ce que cela rapporte ou coûte à un utilisateur d'agir en fonction de la prévision. Ceci demeure donc parfaitement mesurable, même pour une prévision probabiliste isolée. Les scores de qualité couramment utilisés pour ces prévisions probabilistes, lorsqu'ils sont basés sur la composition de Scores Individuels sont donc nécessairement des scores d'utilité. Ils font référence à un certain type d'utilisateur, même si cela n'est pas dit de façon explicite.

3.3.2 Transformation en prévision déterministe

Manipuler des prévisions probabilistes peut sembler contraignant. C'est pourquoi une première façon d'utiliser, d'émettre et donc ensuite d'évaluer de telles prévisions consiste à les transformer en prévisions déterministes. Une fois cela fait, il est possible d'utiliser toute la panoplie des scores de vérification qui s'appliquent à ces dernières. Une telle transformation permet par ailleurs de comparer des systèmes de prévisions probabilistes avec des systèmes déterministes. Ce faisant, on pénalise néanmoins le système probabiliste car une partie de l'information contenue dans la prévision originelle est alors occultée voire perdue.

Probabilité de coupure

Dans le cas de variables non ordonnées (type de précipitation par exemple), on se contente généralement d'annoncer la catégorie possédant la plus forte probabilité prévue comme prévision déterministe.

Si l'on considère une variable ordonnée toutefois, la première étape de la transformation d'une prévision probabiliste en prévision déterministe consiste à se fixer une *probabilité de coupure* P_c :

- si la prévision probabiliste se présente sous la forme d'une fonction de distribution continue $F^{prevue}(x)$, alors la prévision déterministe associée, notée $x^{*,det}$ est la valeur qui vérifie la relation :

$$F^{prevue}(x^{*,det}) = P_c$$

- si la prévision est exprimée en catégories ordonnées entre elles, alors la catégorie prévue est la catégorie G_{det} qui vérifie la relation :

$$\sum_{k=1}^{G_{det}-1} p_k < P_c \leq \sum_{k=1}^{G_{det}} p_k$$

où p_k est la probabilité associée à la catégorie k par la prévision. Le danger de tels découpages en catégories est que les résultats obtenus sont fortement dépendants de la façon dont ces catégories sont définies.

Le choix de la probabilité de coupure est crucial. En effet, si elle est choisie trop élevée, le risque de fausse alerte devient important, alors que si elle est trop faible, la capacité de détection s'amenuise. Si l'on dispose de la fonction de coût de l'utilisateur final, on pourra déterminer la valeur optimale de P_c en considérant que c'est celle qui minimise l'espérance du coût qu'aura à subir cet utilisateur s'il se prépare pour $x^{*,det}$ (cf. 3.3.5.1). On mesure alors plutôt une utilité de la prévision.

Si cette fonction de coût nous est inconnue, il est nécessaire de disposer d'un échantillon de prévisions et d'observations. La probabilité de coupure sera alors choisie de telle manière qu'elle maximise, sur cet échantillon, le score de vérification choisi. Quand on s'intéresse à des prévisions en catégories, il est courant de considérer le score d'Hanssen-Kuipers TSS ou le Biais B .

Diagramme ROC

En considérant successivement différentes valeurs de P_c , il est possible de proposer un score de vérification de la prévision probabiliste basé sur l'ensemble des scores de qualité ainsi obtenus (*i.e.* en faisant varier P_c).

C'est ce principe qui est utilisé dans la création des diagrammes ROC (Receiver Operating Characteristic). Cet outil a principalement été employé en médecine et est utilisé depuis peu en météorologie. Il y a donc au sein de la communauté météorologique un certain flou quant à sa définition exacte. Nous adoptons ici la définition couramment utilisée en médecine et reprise par Stanski *et al.* (1989), cité par Golding (2000).

Ce diagramme s'obtient en traçant la Probabilité de Détection (POD) en fonction du Taux de Fausse Détection ($1 - SPE$) pour différentes valeurs de probabilité de coupure P_c . Si la prévision est initialement continue, il faut de plus préalablement se fixer un seuil X_s afin d'être capable d'établir un tableau de contingence (cf. 3.2.2.1). A chaque valeur du seuil X_s correspond une courbe ROC différente.

On obtient ainsi une droite qui se confond avec la première bissectrice dans le cas d'une prévision respectant la fréquence climatologique d'alerte mais émise complètement au hasard. Plus la prévision est performante, plus cette courbe est proche de l'angle supérieur gauche ($POD = 1$ et $1 - SPE = 0$), comme illustré par la figure 3.9. On peut remarquer que dans un tel diagramme, les lignes d'isovaleurs de TSS sont les parallèles à la première bissectrice, du fait que $TSS = POD - (1 - SPE)$.

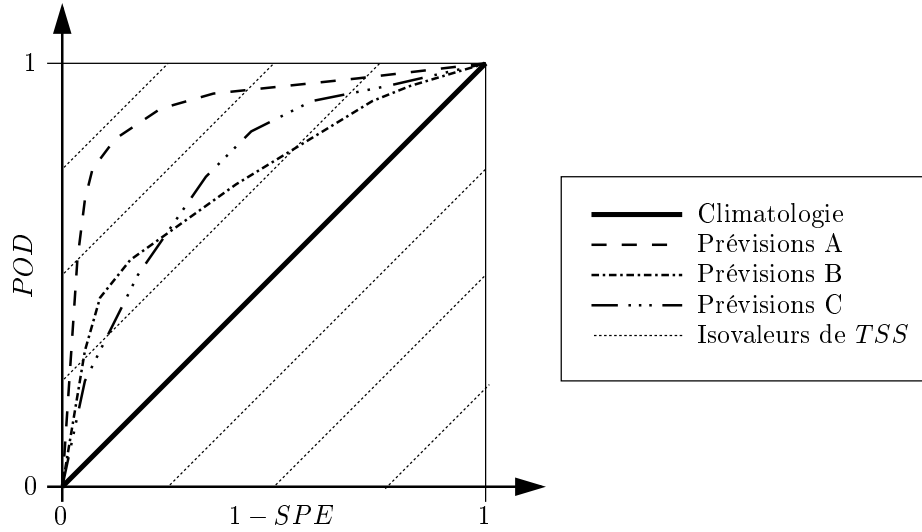


FIG. 3.9: Exemple de courbe ROC. On constate la supériorité de la prévision A sur les prévisions B et C. Les prévisions B et C présentent sensiblement la même aire sous leurs courbes ROC et obtiennent donc le même score. Le diagramme montre de plus que la prévision C a tendance à avoir globalement une meilleure Probabilité de Détection, alors que la prévision B présente un Taux de Fausse Détection plus intéressant.

L'aire située sous la courbe ROC est souvent utilisée comme un score, qui vaut 0,5 pour une prévision faite au hasard en respectant la climatologie de l'événement, et 1 pour une prévision parfaite. Ce score permet de quantifier le caractère globalement discriminant d'une prévision entre deux événements.

Bien qu'associé à une prévision probabiliste, le diagramme ROC analyse celle-ci en la transformant en une succession de prévisions déterministes. Par contre, nous allons maintenant considérer des critères d'évaluation qui préservent d'emblée le caractère probabiliste. Nous présenterons d'abord des critères proposés par d'autres auteurs (*le "prêt-à-porter"*), puis nos propres propositions (*le sur-mesure*).

3.3.3 Le prêt-à-porter (quelques critères classiques)

3.3.3.1 Probability Score (PS)

Le Probability Score a été proposé par Brier (1950). Il s'applique à des prévisions probabilistes par catégories et s'exprime par :

$$PS = \frac{1}{N} \sum_{i=1}^N \left[\sum_{k=1}^G (p_{i,k} - \delta_{i,k})^2 \right] \quad (3.14)$$

où :	N est le nombre d'éléments de l'échantillon de prévisions sur lequel est calculé le score,
	G est le nombre de catégories mutuellement exclusives et exhaustives servant de support à la prévision,
	$p_{i,k}$ est la probabilité associée à la catégorie k par la prévision pour l'élément i de l'échantillon,
	$\delta_{i,k}$ est une fonction qui prend la valeur 1 si l'observation i appartient à la catégorie k et 0 sinon.

Le Probability Score est égal à 0 dans le cas de prévisions parfaites et à 2 dans le cas des plus mauvaises prévisions possibles (prévision d'une probabilité de 100% dans une catégorie autre que la catégorie observée).

On peut montrer que dans le cas général, pour un nombre quelconque de catégories mutuellement exclusives et exhaustives, le PS est équivalent à une mesure de justesse (Epstein et Murphy 1965). Néanmoins, dans le cas de prévision OUI/NON (deux catégories), le PS peut être décomposé en plusieurs termes (Sanders 1963; Murphy 1973), mesurant le biais à grande échelle et la finesse. Bien qu'ancien, ce score reste encore couramment utilisé (voir Atger 1999 par exemple), notamment dans les cas à 2 catégories.

3.3.3.2 Ranked Probability Score (RPS)

Le principal inconvénient du Probability Score est qu'il ne tient pas compte du fait que les catégories utilisées peuvent être ordonnées entre elles. Ainsi, si l'on considère la prévision présentée par le tableau 3.8, il apparaît que celle-ci se voit attribuer le même score $PS = 1.22$, que la valeur observée soit de 0 mm (catégorie 1) ou de 4 mm (catégorie 2). Or, il est clair pour tout un chacun que cette prévision devrait être moins pénalisée dans le deuxième cas, la valeur observée se trouvant alors moins éloignée du mode de la prévision (catégorie 3) que dans le premier cas.

	Classes de précipitation R (mm)			
	$R = 0$	$0 < R \leq 5$	$5 < R \leq 20$	$20 < R$
Probabilité d'appartenance à la classe (%) :	10	10	60	20

TAB. 3.8: Exemple de prévision probabiliste de précipitation par classe de pluie.

Le Ranked Probability Score, introduit par Epstein (1969) puis reformulé par Murphy (1971) sous la forme actuellement utilisée, permet de répondre à ce problème. Il s'exprime, avec les notations définies précédemment, de la manière suivante :

$$RPS = \frac{1}{N} \sum_{i=1}^N \left[\sum_{k=1}^G \left(\sum_{l=1}^k p_{i,l} - \sum_{l=1}^k \delta_{i,l} \right)^2 \right] \quad (3.15)$$

Le RPS varie de 0, pour une prévision parfaite, à $G - 1$ dans le cas des plus mauvaises prévisions possibles : prévision d'une probabilité de 100% dans la catégorie 1 (resp. G) alors que la catégorie observée est G (resp. 1).

Le principe de ce score consiste à comparer non pas les probabilités prévues et observées classe par classe, comme le fait le PS , mais plutôt à comparer les probabilités au non-dépassement prévues et observées. De cette manière le RPS tient compte du caractère ordonné des catégories et favorise les prévisions dont le centre de gravité est le plus proche de la valeur observée. Ainsi, si l'on reprend l'exemple de la prévision présentée par le tableau 3.8, on obtient les scores $RPS = 1.59$ si la valeur observée est de 0 mm et $RPS = 0.69$ si elle est de 4 mm.

Il est évident que ce score n'a aucun sens dans le cas de catégories non-ordonnées, et dans ce cas le PS reste plus significatif. En outre, si l'on ne considère que 2 catégories, le RPS n'apporte bien sûr rien de plus que le PS et l'on a : $RPS = \frac{1}{2}PS$.

3.3.3.3 Continuous Ranked Probability Score (CRPS)

Le Ranked Probability Score, précédemment présenté, n'est pas lui non plus exempt d'inconvénients :

- la plage de valeurs admissibles dépend du nombre de catégories considérées,
- la même prévision, exprimée selon divers découpages de catégories, se verra attribuer des scores RPS différents. Dans le contexte d'une évaluation, des personnes peu scrupuleuses peuvent même imaginer d'optimiser le découpage pour maximiser le RPS ¹ !

Pour s'affranchir de ces problèmes, il suffit d'abandonner l'idée d'une variable sagement rangée dans un nombre prédéterminé de catégories. La variable de la prévision est alors autorisée à se mouvoir sur un support continu.

Pour ce faire, considérons les G catégories formant une partition de notre « réalité » : $([b_l; b_{l+1}[)_{l=1,G}$. Si l'on effectue un passage à la limite sur $G \rightarrow +\infty$, *i.e.* on considère un nombre de catégories de plus en plus important, alors la probabilité p_l associée à la catégorie l devient égale à une densité de probabilité $f(b_l)$:

$$\lim_{G \rightarrow +\infty} p_l = f(b_l) \quad (3.16)$$

Selon le même raisonnement, la somme des probabilités de toutes les catégories inférieures à la catégorie k devient égale à la valeur d'une fonction de répartition :

$$\lim_{G \rightarrow +\infty} \sum_{l=1}^k p_l = \int_{-\infty}^{b_k} f(x).dx = F(b_k) \quad (3.17)$$

¹L'auteur tient à préciser qu'il est outré devant de telles pratiques que la morale réprouve à juste titre !

L'expression du *RPS* selon un support continu conduit donc à l'expression du Continuous Ranked Probability Score *CRPS* (Brown 1974 ; Matheson et Winkler 1976 ; Bouttier 1994) :

$$CRPS = \frac{1}{N} \sum_{i=1}^N \left(\int_{-\infty}^{+\infty} [F_i(x) - H_{x_i^0}(x)]^2 dx \right) \quad (3.18)$$

où :

- en plus des notations précédentes :
- x_i^0 est la valeur observée pour l'élément i de l'échantillon de prévisions,
- F_i est la fonction de répartition prévue pour l'élément i de l'échantillon de prévisions,
- $H_{x_i^0}$ est une fonction telle que : $H_{x_i^0}(x) = \begin{cases} 0 & \text{si } x < x_i^0 \\ 1 & \text{si } x \geq x_i^0 \end{cases}$

Le *CRPS* s'interprète aisément graphiquement à l'aide des fonctions de répartition prévues (F_i) et observées ($H_{x_i^0}$). Il correspond à l'aire comprise entre F_i^2 et $H_{x_i^0}$ pour tout x inférieur à x_i^0 et à l'aire comprise entre $(1 - F_i)^2$ et $H_{x_i^0}$ pour tout x supérieur (surface hachurée sur la figure 3.10).

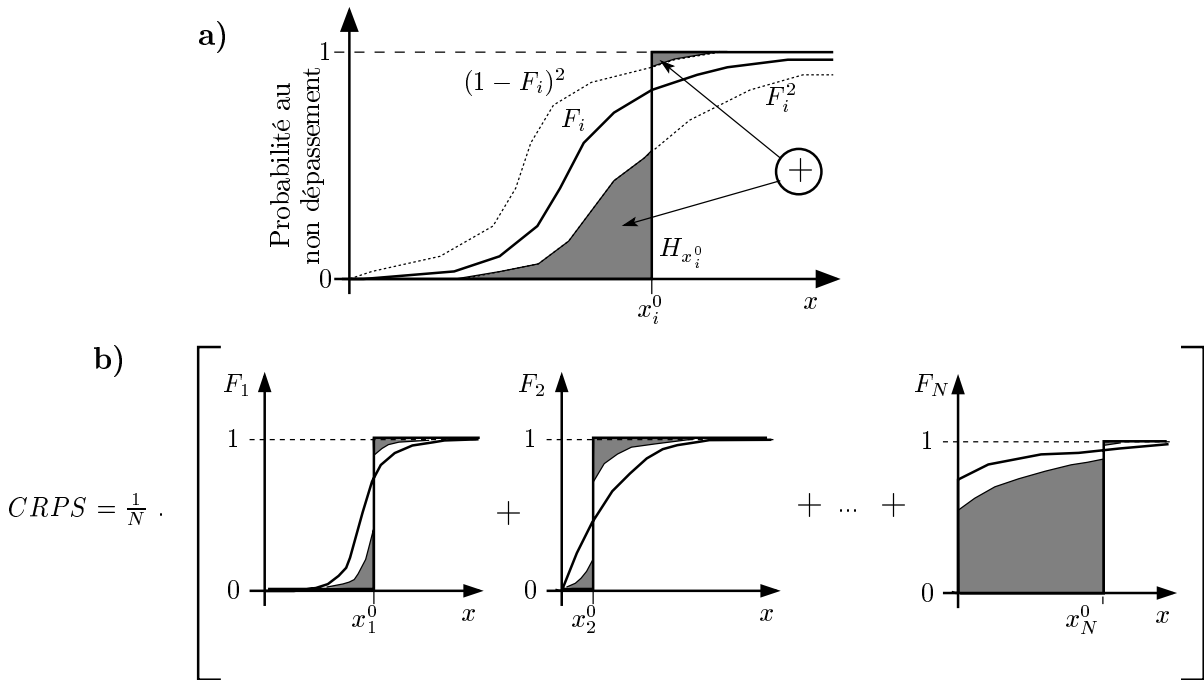


FIG. 3.10: Interprétation graphique du CRPS : (a) pour un seul couple prévision/observation (Score Individuel), (b) pour un échantillon de N couples prévision/observation (Score Global).

Hersbach (2000) a montré que le *CRPS* est un score non-biaisé et a proposé une décomposition qui permet de faire apparaître la part du score due à la présence de biais à grande échelle.

Une autre décomposition a été proposée par Candille (2003). Elle permet d'exprimer le *CRPS* sous la forme de termes de *fiabilité* (qui correspond ici à un biais à petite échelle), de *résolution* (*i.e.* finesse) et d'*incertitude* (partie du score due à la climatologie des observations sur l'échantillon). Cette décomposition est très attrayante théoriquement parlant, mais sa mise en œuvre

pratique, comme cela a été souligné par son auteur, est extrêmement délicate. Elle est notamment fortement limitée par la taille N des échantillons considérés, du fait de la tentative de mesure du biais à petite échelle.

Comme nous l'avons vu dans la partie 3.1.3.4, la mesure du biais à petite échelle, bien que très importante, ne nous est pas vraiment accessible. C'est pourquoi on se contente souvent du biais à grande échelle. Néanmoins, un défaut de biais à grande échelle peut-être corrigé a posteriori et c'est pourquoi nous pensons que la présence d'un terme de biais n'est pas indispensable dans la décomposition d'un score tel que le *CRPS*. Nous proposerons donc plus loin (cf. 3.3.5), en considérant le *CRPS* comme un score d'utilité, une nouvelle décomposition de ce score en terme de *finesse* et de *justesse*, qui présente l'avantage d'être facilement mise en application.

3.3.4 Le prêt-à-porter avec retouches (utilisation de la Vraisemblance)

Au sein de ce sous-chapitre, nous tentons la mise en pratique d'une idée a priori alléchante : l'adaptation de la notion de vraisemblance à une utilisation comme score de vérification. Cet essai ne s'est malheureusement pas révélé applicable à notre contexte. La lecture de cette partie n'est donc pas indispensable pour le lecteur dont le temps est compté.

3.3.4.1 Définition

Lorsque l'on veut associer une loi de probabilité à une variable aléatoire pour laquelle on dispose d'un certain nombre de réalisations, on utilise couramment la notion de *vraisemblance*. Dans son utilisation la plus courante, on procède ainsi : à partir d'un échantillon de N réalisations indépendantes $(x_i)_{i=1,N}$, on choisit tout d'abord *a priori* un certain nombre de lois de probabilités possibles \mathcal{L}_j . Ensuite, afin de déterminer selon laquelle de ces lois il est le plus probable que la variable aléatoire soit régie, on calcule, pour chacune d'elles, la vraisemblance V_j sur cet échantillon :

$$V_j = \prod_{i=1}^N f_j(x_i) \quad (3.19)$$

où f_j est la densité de probabilité de la loi \mathcal{L}_j . La loi à laquelle est associée la plus forte valeur de vraisemblance est considérée comme étant la plus probable pour décrire cet échantillon.

Or, il est possible d'étendre cette notion en considérant non plus N réalisations d'une seule variable aléatoire, mais plutôt une série de mono-réalisations de N variables aléatoires, comme proposé par Obled (note interne LTHE, 1997). Si on dispose de deux (ou plus) méthodes de prévisions probabilistes $M1$ et $M2$, chacune d'elles fournit, pour une même série de N prévisions, N lois de probabilité (une pour chaque prévision i) : F_i^{M1} et F_i^{M2} , soit exprimé en densité : f_i^{M1} et f_i^{M2} . On peut alors rechercher quel est l'ensemble de N lois de probabilités qui décrit le mieux ces N événements. Dans le cadre de la vérification d'une série de prévisions, les variables aléatoires et les lois de probabilité sont respectivement les valeurs x_i^0 mesurées et les prévisions probabilistes émises. On calcule alors la vraisemblance V_{M1} (resp. V_{M2}) d'une série de prévisions issues d'un système $M1$ (resp. $M2$) sur un échantillon de N observations x_i^0 :

$$V_{M1} = \prod_{i=1}^N f_i^{M1}(x_i^0) \quad (3.20)$$

où f_i^{M1} est la densité de probabilité de la prévision issue du système $M1$ pour le jour i .

Ce faisant, chaque méthode de prévision probabiliste affecte à la série d'observations une certaine vraisemblance. La méthode qui associe la plus grande vraisemblance aux observations réalisées est celle qui les considère comme les plus probables. C'est évidemment cette méthode qui doit être retenue comme étant la plus performante, puisque les observations se sont effectivement réalisées ! A la limite, une méthode de prévision « parfaite » dira que seules les observations réalisées étaient possibles (probabilité associée de 1) et attribuera à l'échantillon observé une vraisemblance maximale ($+\infty$). L'extrême inverse est représenté par une méthode qui verrait a priori toutes les observations réalisées comme impossibles (probabilité exactement nulle) : la vraisemblance est alors nulle.

Dans la pratique, pour des raisons de calcul numérique, il peut être intéressant de travailler avec $\log(V)/N$. Ceci évite d'atteindre des valeurs infiniment petites difficilement gérables. Une prévision parfaite atteint alors toujours un score de $+\infty$, les moins efficaces obtenant un score de $-\infty$.

3.3.4.2 Attributs mesurés

La vraisemblance mesure tout à la fois la justesse et la finesse :

- La justesse : la valeur x_i qui maximise $f_i^{M1}(x_i)$, à savoir le mode de la distribution prévue, est par définition la valeur qui est considérée comme la plus probable par la prévision. La vraisemblance mesure donc un écart entre la réalité et cette valeur.
- La finesse : la loi de probabilité qui maximise $f_i^{M1}(x_i)$ correspond à la loi de finesse parfaite qui donne l'événement x_i comme certain : f_i^M est alors un Dirac en x_i^0 . La vraisemblance mesure donc aussi un écart à cette loi de finesse parfaite. Il convient toutefois de noter que cette mesure de la finesse n'est pas indépendante de la valeur x_i observée *a posteriori* : la

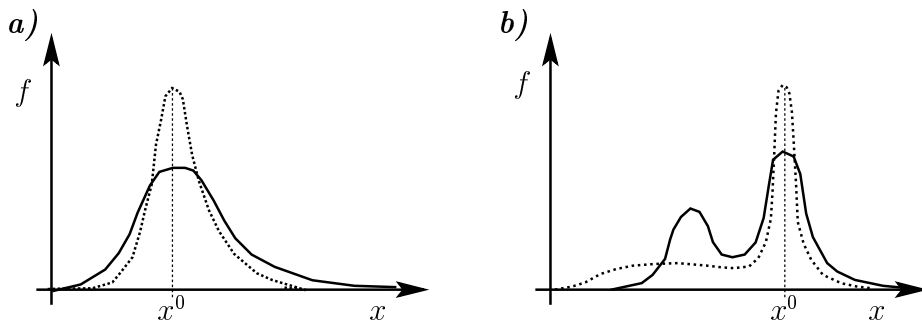


FIG. 3.11: Exemples de densités de probabilités prévues de justesses identiques. **a)** la prévision en pointillés affiche des finesesses locales et globales supérieures à celles de la prévision en trait plein, **b)** la prévision en pointillés affiche une finesse locale supérieure à celle de la prévision en trait plein, mais une finesse globale inférieure.

finesse considérée ici est une finesse locale. Etant donné que par définition $\int_{-\infty}^{+\infty} f_i(t).dt = 1$, cette finesse locale est assimilable à la finesse globale pour des lois de probabilité classiques, telles que la loi normale (fig 3.11 a). Cependant, pour des lois plus exotiques (lois bimodales par exemple), cela peut ne pas être toujours vrai (fig 3.11 b).

En outre, la vraisemblance est un score non-biaisé. Considérons ainsi un échantillon de N observations $(x_i^0)_{i=1,N}$, de densité de probabilité empirique marginale f_{Clim} , sur lequel a été émise à chaque fois la même prévision de densité de probabilité f^{M1} (c'est-à-dire que la prévision probabiliste suit toujours la même loi de distribution). On a alors :

$$V_{f^{M1}} \geq V_{f_{Clim}} \Leftrightarrow \prod_{i=1}^N f^{M1}(x_i^0) \geq \prod_{i=1}^N f_{Clim}(x_i^0) \quad (3.21)$$

$$\Leftrightarrow \frac{1}{N} \sum_{i=1}^N \log(f^{M1}(x_i^0)) \geq \frac{1}{N} \sum_{i=1}^N \log(f_{Clim}(x_i^0)) \quad (3.22)$$

Soit $\{y_j\}_{j=1,M}$ l'ensemble des M valeurs prises par les N observations $x_{i=1,N}^0$. Par définition de la fonction f_{Clim} , l'équation 3.22 peut s'écrire, en sommant sur les y_j :

$$V_{f^{M1}} \geq V_{f_{Clim}} \Leftrightarrow \sum_{j=1}^M f_{Clim}(y_j) \cdot \log(f^{M1}(y_j)) \geq \sum_{j=1}^M f_{Clim}(y_j) \cdot \log(f_{Clim}(y_j)) \quad (3.23)$$

$$\Leftrightarrow \sum_{j=1}^M f_{Clim}(y_j) \cdot \log\left(\frac{f^{M1}(y_j)}{f_{Clim}(y_j)}\right) \geq 0 \quad (3.24)$$

$$\Leftrightarrow \sum_{j=1}^M f_{Clim}(y_j) \cdot \log\left(\frac{f_{Clim}(y_j)}{f^{M1}(y_j)}\right) \leq 0 \quad (3.25)$$

Le premier terme du membre de droite de la relation 3.25 est ce que l'on appelle la divergence de Kullback-Leibler de f^{M1} à f_{Clim} , que l'on note $K(f^{M1}|f_{Clim})$, d'où :

$$V_{f^{M1}} \geq V_{f_{Clim}} \Leftrightarrow K(f^{M1}|f_{Clim}) \leq 0 \quad (3.26)$$

Or, les mathématiciens nous disent que, quelles que soient les densités de probabilités f^{M1} et f_{Clim} , $K(f^{M1}|f_{Clim})$ est toujours positive et devient nulle si et seulement si $f^{M1} = f_{Clim}$ (Pham 2001, page 25). On en déduit donc que :

$$V_{f^{M1}} \geq V_{f_{Clim}} \Leftrightarrow K(f^{M1}|f_{Clim}) = 0 \quad (3.27)$$

$$\Leftrightarrow f^{M1} = f_{Clim} \quad (3.28)$$

ce qui signifie que la prévision qui minimise la vraisemblance sur une série de prévisions identiques, correspond à la climatologie empirique sur l'échantillon considéré. La vraisemblance est donc bien un score non-biaisé.

3.3.4.3 Inconvénients

Malgré des attributs fort attrayants, la vraisemblance n'est pas un score parfait. En effet, son principal inconvénient est qu'il requiert l'expression de la densité de probabilité de la prévision. Or, les prévisions probabilistes sont généralement obtenues en générant un échantillon de réalisations possibles de l'événement : on dispose ainsi d'une fonction de répartition qui n'est pas exprimée de façon analytique. Pour en déduire une densité de probabilité, il faut lui faire subir d'abord un ajustement, puis une dérivation, ce qui peut se heurter à quatre difficultés :

- ① Il faut se fixer *a priori* la forme que doit avoir la densité de probabilité, c'est-à-dire préciser quel type de loi régit l'événement. En ce qui concerne les distributions de précipitations journalières, on considère en général une loi Gamma incomplète. Cependant, lorsque l'on veut calculer la vraisemblance de façon systématique sur une série de prévisions, il est peu pratique de s'assurer pour chaque prévision que l'ajustement de la loi de probabilité est réaliste. Il existe bien des méthodes statistiques pour rejeter des ajustements, mais il n'est pas réaliste d'aller choisir « à la main » une loi de probabilité différente à chaque fois que la loi choisie par défaut est rejetée. On risque donc de calculer un score sur des hypothèses éventuellement erronées.
- ② Dès que l'on se trouve dans les queues de distributions (ce qui finit malheureusement toujours par arriver, aussi efficace que soit la prévision), les valeurs de densités de probabilité peuvent devenir extrêmement faibles. Ceci peut poser des problèmes numériques non négligeables.
- ③ La présence de points d'accumulation peut elle aussi être source d'ennuis. Ainsi, dans le cas des prévisions de précipitation, il est très fréquent que la probabilité d'avoir une pluie exactement nulle (*i.e.* absence de précipitation) ne soit pas nulle : $f(0)$ n'est pas défini, alors que $F(0)$ existe. Il faut alors se méfier et ne pas remplacer $f(0)$ par $F(0)$ dans le calcul de la vraisemblance : ce ne sont pas des quantités homogènes. Pour s'en persuader, il suffit de se souvenir que F est borné entre 0 et 1, alors que f ne connaît pas de limite supérieure. Pour s'en sortir, la meilleure façon est de transformer les prévisions continues en prévisions par catégories. On définit ainsi N_c classes distinctes $([b_k; b_{k+1}[)_{k=1, N_c}$ dont l'union décrit le support de la variable cible. La vraisemblance est alors définie dans ce cas par (avec les notations définies précédemment) :

$$V_{M1} = \prod_{i=1}^N \left(\sum_{k=1}^{N_c} (F_{i,k}^{M1} \cdot \delta_{i,k}) \right) \quad (3.29)$$

$$\text{avec : } F_{i,k}^{M1} = \int_{b_k}^{b_{k+1}} f_i^{M1}(t).dt \quad \text{et} \quad \begin{cases} \delta_{i,k} = 1 & \text{si } x_i^0 \in [b_k; b_{k+1}[\\ \delta_{i,k} = 0 & \text{sinon} \end{cases} \quad (3.30)$$

Ce faisant, on dégrade la prévision en se ramenant à une prévision par catégories et l'on se trouve confronté au problème de la définition des bornes des classes (idéalement équiprobable selon la climatologie), ce qui ne constitue pas non plus une sinécure...

④ Enfin, dans une quarantaine de cas sur plusieurs milliers testés, nous avons été confronté au cas de figure suivant :

- une prévision probabiliste catégorique de pluie nulle est émise (certaines situations anti-cycloniques),
- alors que la pluie observée est finalement non nulle (de 0,5 à 4 mm) du fait de l'arrivée un peu plus rapide que prévu d'une perturbation, ou de la disparition d'un blocage.

Dans ce cas, la vraisemblance de cette prévision est strictement nulle, et donc la vraisemblance de l'ensemble de la série est elle aussi nulle, quoi que valent les prévisions pour les autres éléments de la série... Pour éviter ces problèmes, il suffit de retirer ces quelques cas de notre échantillon de prévision, mais cette censure laisse planer le doute d'un biais dans les résultats alors obtenus.

De même que dans le point précédent, ce problème est lié à la présence d'un point d'accumulation dans la fonction de répartition de notre variable. Dans le cas d'une variable telle que la température, sans point d'accumulation, il est en effet extrêmement improbable d'émettre une prévision catégorique.

3.3.4.4 Conclusion

Même si cette approche par vraisemblance a été testée avec un certain succès, elle a été finalement laissée de côté. Les difficultés de mise en œuvre dans le cadre d'une variable présentant un point d'accumulation comme la précipitation sont assez importantes et imposent des censures peu satisfaisantes. Pour d'autres variables ne présentant pas de points d'accumulation, telle la température, il est hautement improbable de faire une prévision catégorique ou présentant elle aussi un point d'accumulation. Ces problèmes ne se posent donc pas et l'approche par vraisemblance reste une possibilité intéressante.

3.3.5 Le Sur-Mesure (de la *Fonction de Coût* au *Coût Effectif Sociétal*)

Tous les scores présentés précédemment n'expriment pas de façon explicite à quels types d'utilisateurs ils font référence. Ils définissent une utilité *ex nihilo* et c'est à l'utilisateur de décider s'ils sont adaptés à ses objectifs. Cependant, si l'utilisateur est capable de quantifier ces derniers, il peut sembler plus intéressant de construire directement un score qui réponde à ses attentes.

Dans cette partie, nous généralisons la démarche déjà adoptée par Epstein (1969). Celui-ci s'était placé dans un contexte de prévisions discrètes et n'avait considéré qu'une seule catégorie d'utilisateurs (*i.e.* la fonction de coût était fixée). Avec notre généralisation, nous obtiendrons des résultats similaires à ceux obtenus par Matheson et Winkler (1976), mais l'approche adoptée ici nous permettra :

- d'une part, d'explicitier clairement la démarche utilisée selon le point de vue d'un utilisateur potentiel,
- d'autre part, de proposer une décomposition des scores d'utilité ainsi obtenus en terme de finesse et de justesse de la prévision,
- et enfin, de démontrer que le CRPS est bien un score d'utilité.

3.3.5.1 Une méthode générale de construction d'un score d'utilité

Démarche adoptée

Après avoir défini ce qu'est une fonction de coût (étape ①), nous allons construire notre score d'utilité en deux étapes :

- ① Dans la première étape, nous allons considérer un unique utilisateur muni de sa fonction de coût. Nous allons voir que celui-ci est capable, à partir d'une prévision probabiliste :
- de choisir la valeur optimale pour laquelle il doit se préparer,
 - et d'évaluer combien ce choix lui a coûté a posteriori, une fois l'observation réalisée.

Nous allons ainsi définir la notion de **coût effectif** qui permet à cet utilisateur de décider entre deux méthodes de prévision laquelle lui est le mieux adaptée. Ce coût effectif forme malheureusement un score d'utilité peu généralisable.

- ② Lors de la deuxième étape, nous considérerons, non plus un unique utilisateur, mais un ensemble fini d'utilisateurs ayant chacun leur fonction de coût respective. Cette fois-ci, nous allons :
- choisir, pour chaque utilisateur, la valeur optimale pour laquelle il doit se préparer,
 - et évaluer, globalement pour l'ensemble des utilisateurs, ce que ces choix leur ont coûté a posteriori.

Nous définirons ainsi le **coût effectif moyen** pour cet ensemble d'utilisateurs. Ensuite, en augmentant significativement leur nombre, pour qu'ils deviennent représentatifs de la « société », nous arriverons à la notion de **coût effectif sociétal**, qui permet d'identifier la méthode de prévision la mieux adaptée au plus grand nombre.

Ensuite, l'étape ③ va montrer que ce coût effectif sociétal peut être décomposé en terme de finesse et de justesse. Cette propriété nous amenera tout naturellement à considérer le coût effectif sociétal comme un score d'utilité tout à fait satisfaisant.

Enfin, afin d'éviter au lecteur de se noyer dans les tumultueux développements mathématiques qui suivent, nous nous plaçons, lors de ces premières étapes, dans le cadre d'un unique couple prévision/observation et non plus dans celui d'une série de couples prévision/observation. Ce faisant, nous sommes amenés à considérer seulement des Scores Individuels : les indices i décrivant les éléments de l'échantillon de prévisions considéré peuvent donc être omis des équations. Pour obtenir des Scores Globaux, il suffira d'appliquer une fonction de composition h aux Scores Individuels obtenus (cf. 3.1.2.2), comme nous le verrons dans l'étape ④.

① La fonction de coût

La construction d'un score d'utilité requiert tout d'abord la définition de la *fonction de coût* \mathcal{C} de l'utilisateur.

Soit x la grandeur concernée par la prévision. Le coût correspond à la perte financière subie par l'utilisateur¹ :

¹Dans le contexte de l'utilisation de prévisions météorologiques pour la gestion des risques naturels, on ne considère que l'éventualité de pertes : il n'y a rien à gagner avec les risques naturels, mais beaucoup à perdre!

- lorsque, ayant reçu une prévision, il se prépare pour une certaine valeur x^* de x (valeur cible),
- alors que la réalisation s'avère a posteriori être x^0 (la réalisation).

On établit ainsi la fonction de coût $\mathcal{C}(x^*; x^0)$ qui permettra d'attribuer un coût à chaque couple cible/réalisation $(x^*; x^0)$.

Cette fonction de coût est propre à un utilisateur donné et caractérise l'utilisation qu'il fait de la prévision. Elle n'est pas nécessairement la même tous les jours. Ainsi, dans le cas de prévisions de précipitations destinées aux agriculteurs, l'impact d'une bonne prévision n'est pas le même en hiver, alors que les sols se reposent, et à la fin du printemps, lorsque les plantes sont en pleine croissance. Néanmoins, par souci de clarté, nous considérons dans la suite des fonctions de coût indépendantes du temps.

Pour être acceptable, une fonction de coût doit respecter les deux conditions suivantes :

$$\forall x^*, x^0 \quad : \quad \mathcal{C}(x^*; x^0) \geq 0 \quad (3.31)$$

$$\forall x^*, x^0 \quad : \quad \mathcal{C}(x^*; x^0) = 0 \quad \Leftrightarrow \quad x^* = x^0 \quad (3.32)$$

Ces conditions sont peu restrictives. Ainsi, si la relation 3.31 n'est pas respectée, dans le cas de gains par exemple, il suffit de remplacer la fonction de coût \mathcal{C} par $\mathcal{C} - \min \mathcal{C}$, ce qui permet par une pirouette de transformer les gains en pertes¹.

① *Le coût effectif*

Une fois que l'on dispose de cette fonction \mathcal{C} et d'une prévision \mathcal{P} donnant une loi de probabilité de densité $f(x)$, il est possible de calculer le coût que l'on s'attend à payer en espérance si l'on décide de se préparer pour la valeur cible x^* (*i.e.* en moyenne si l'on reproduisait de nombreuses fois l'expérience). Le fait de disposer de la prévision \mathcal{P} permet effectivement de supposer que la valeur x^0 que l'on va observer suit la densité de probabilité $f(x)$. On définit ainsi l'espérance attendue du coût $E^{\mathcal{C}, \mathcal{P}}(x^*)$:

$$E^{\mathcal{C}, \mathcal{P}}(x^*) = \int_{-\infty}^{+\infty} \mathcal{C}(x^*, x^0) \cdot f(x^0) \cdot dx^0 \quad (3.33)$$

qui est une fonction de x^* . En effet, l'utilisateur ne peut plus intervenir ni sur la fonction de coût préétablie, ni sur la prévision qui lui est fournie : sa seule latitude réside dans le choix de la valeur cible x^* .

Dans d'autres contextes (gestion de ressources en eau ou hydroélectricité par exemple), l'utilisation d'une bonne prévision peut évidemment engendrer de réels gains.

¹L'auteur est conscient de l'incongruité financière que cela représente de transformer des gains en pertes. Le lecteur choqué par ces démarches contre-productives est autorisé à travailler non plus avec une fonction de coût, mais avec une fonction de gain. Les raisonnements présentés ici sont parfaitement analogues, les seules différences résidant dans le changement de sens des inégalités et l'inversement des opérateurs min et max. Il est ainsi possible de transformer des pertes en gains, ce qui est certes plus productif, mais demeure tout aussi répréhensible dans les bilans comptables.

La question de l'existence de cette espérance est aisément résolue en se limitant aux prévisions dont les densités de probabilités f sont rigoureusement nulles hors d'un certain intervalle : $\exists X_{lim}/\forall x \in [-X_{lim}; X_{lim}] : f(x) = 0$.

Dans la pratique, l'utilisateur cherche à minimiser ses pertes à partir des seules informations dont il dispose *a priori*. C'est à dire qu'il va se préparer pour la valeur cible optimale¹ x_{opt}^* qui minimise l'espérance attendue du coût $E^{\mathcal{C}, \mathcal{P}}(x^*)$:

$$\min E^{\mathcal{C}, \mathcal{P}}(x^*) = E^{\mathcal{C}, \mathcal{P}}(x_{opt}^*) \quad \text{donc} \quad \left. \frac{\partial E^{\mathcal{C}, \mathcal{P}}(x^*)}{\partial x^*} \right|_{x_{opt}^*} = 0 \quad (3.34)$$

Après la réalisation effective de l'événement prévu, avec pour valeur x_{obs}^0 , l'utilisateur peut faire ses comptes. Il s'est préparé pour x_{opt}^* alors que la réalisation est x_{obs}^0 : le coût effectif de cet événement est donc : $\mathbf{Ceff} = \mathcal{C}(x_{opt}^*; x_{obs}^0)$.

Le coût effectif est souvent utilisé comme mesure de l'utilité d'une prévision isolée. Ce n'est toutefois pas celui que l'on utilisera ici, car il est trop spécifique à cet unique utilisateur, ce qui entraîne deux importantes limitations dans son usage :

- ① Tout d'abord, en vertu des relations 3.32 et 3.31, \mathbf{Ceff} est une mesure de distance entre x_{opt}^* et x_{obs}^0 , ce qui correspond tout à fait à la définition de la mesure de la justesse. \mathbf{Ceff} ne mesure donc absolument pas la finesse des prévisions. Ainsi, deux prévisions distinctes, émises par deux méthodes différentes pour la même journée, auront la même utilité au sens du coût effectif à partir du moment où elles conduisent à la même valeur de x_{opt}^* , quelles que soient leurs différences par ailleurs.
- ② Ensuite, ce score postule que l'utilisateur est toujours parfaitement rationnel et se prépare à chaque fois pour la valeur optimale x_{opt}^* . Or, cela peut ne pas être toujours le cas. Que vaut alors la prévision si l'utilisateur n'est pas aussi rationnel que prévu ? Ainsi, prenons l'exemple d'une fonction de coût pour laquelle il est beaucoup plus pénalisant de sous-estimer l'événement que de le surestimer :

$$\mathcal{C}(x^*; x^0) = \begin{cases} 10.(x^0 - x^*) & \text{si } x^0 > x^* \\ (x^* - x^0) & \text{si } x^0 \leq x^* \end{cases} \quad (3.35)$$

Dans ce cas, l'espérance attendue du coût est minimisée par le quantile 91% ($Q_{91\%}$). La meilleure prévision, au sens du coût effectif, est donc celle pour laquelle le $Q_{91\%}$ correspond à la valeur observée. Toutefois, par mesure de sécurité, on préférerait que celle-ci soit associée à un quantile moins élevé.

En effet, si l'on cherche à optimiser une méthode de prévision pour cet utilisateur en se basant sur le coût effectif, nous allons privilégier ces prévisions : en général, il faudra considérer le $Q_{91\%}$ pour avoir une bonne estimation de la valeur qui sera effectivement

¹L'espérance attendue du coût n'admet pas forcément un unique minimum, mais par contre, du fait de la condition 3.31, celui-ci est toujours défini.

observée. Tant que l'utilisateur utilise ces prévisions de façon rationnelle (*i.e.* en cherchant à minimiser l'espérance du coût), il s'attachera effectivement à la valeur du $Q_{91\%}$ et les prévisions lui seront donc utiles. Mais, le temps passant et la routine s'installant, l'utilisateur peut oublier sa fonction de coût, connaître un changement d'opérateur (le nouveau n'étant pas au fait de l'exacte définition de la fonction de coût), ou alors prêter ses prévisions à un autre utilisateur... Dans tous ces cas, l'utilisation de la prévision ne va plus se faire de façon rationnelle : les utilisateurs vont spontanément se focaliser sur des quantiles médians de la prévision et s'écarter du $Q_{91\%}$ qui n'aura plus pour eux de signification particulière. Leurs prévisions vont alors systématiquement sous-évaluer les valeurs effectivement observées, ce qui, du fait de leur fonction de coût, leur coûtera très cher !

Ce score conduit donc à des prévisions trop spécifiques à une fonction de coût donnée : si l'utilisateur ne tient pas compte très exactement de sa fonction de coût lors de l'utilisation de la prévision, il risque gros.

② Le coût effectif sociétal

Si notre démarche, jusqu'à présent, suffit pour un seul utilisateur qui connaît bien ses objectifs et sa fonction de coût (avec les réserves émises précédemment), il demeure néanmoins rare qu'une prévision ne soit émise que pour un seul utilisateur. Nous allons donc construire un score d'utilité qui considère non plus un unique utilisateur, mais plutôt un ensemble d'utilisateurs.

Chacun de ces utilisateurs j possède sa propre fonction de coût \mathcal{C}_j , indépendamment les uns des autres. À partir d'une prévision \mathcal{P} commune pour tous, chacun est donc en mesure de calculer le coût effectif \mathbf{Ceff}_j qui lui est propre :

$$\mathbf{Ceff}_j = \mathcal{C}_j(x_{opt,j}^*; x_{obs}^0) \quad \text{où} \quad x_{opt,j}^* / \min E^{\mathcal{C}_j, \mathcal{P}}(x^*) = E^{\mathcal{C}_j, \mathcal{P}}(x_{opt,j}^*) \quad (3.36)$$

Nous avons donc un coût effectif moyen $\overline{\mathbf{Ceff}}$ sur l'ensemble des M utilisateurs :

$$\overline{\mathbf{Ceff}} = \frac{1}{M} \sum_{j=1}^M \mathbf{Ceff}_j \quad (3.37)$$

$$= \frac{1}{M} \sum_{j=1}^M \mathcal{C}_j(x_{opt,j}^*; x_{obs}^0) \quad (3.38)$$

Certains utilisateurs, bien que ne possédant pas nécessairement la même fonction de coût, peuvent avoir la même valeur cible optimale x_{opt}^* . Ainsi, l'ensemble des valeurs cibles optimales des M utilisateurs $(x_{opt,j}^*)_{j=1,M}$ prend ses valeurs dans l'ensemble $(x_k^*)_{k=1,K}$, avec $K \leq M$. Il est alors intéressant de considérer les M_k utilisateurs dont la valeur cible optimale est x_k^* comme un groupe homogène ayant une seule fonction de coût. Nous définissons ainsi la fonction de coût¹ sociétale $\mathcal{C}_{soc}^{\mathcal{P}}$ qui correspond à la moyenne des fonctions de coût individuelles au sein de ce groupe pour une valeur cible fixée à $x^* = x_k^*$:

¹Il est intéressant de remarquer que la fonction de coût sociétale ainsi définie remplit effectivement les conditions 3.32 et 3.31. C'est donc bien une fonction de coût.

$$\mathcal{C}_{soc}^{\mathcal{P}}(x_k^*; x^0) = \frac{1}{M_k} \sum_{j / x_{opt,j}^* = x_k^*} \mathcal{C}_j(x_k^*; x^0) \quad (3.39)$$

La moyenne $\overline{\mathbf{Ceff}}_k$ des coûts effectifs de ces M_k utilisateurs est donc naturellement donnée par :

$$\overline{\mathbf{Ceff}}_k = \mathcal{C}_{soc}^{\mathcal{P}}(x_k^*; x_{obs}^0) \quad (3.40)$$

L'équation 3.38, donnant le coût effectif moyen pour tous les utilisateurs, peut donc se réécrire en sommant sur les valeurs possibles x_k^* au lieu de le faire directement sur les M utilisateurs :

$$\overline{\mathbf{Ceff}} = \frac{1}{M} \sum_{k=1}^K M_k \cdot \overline{\mathbf{Ceff}}_k \quad (3.41)$$

$$= \frac{1}{M} \sum_{k=1}^K M_k \cdot \mathcal{C}_{soc}^{\mathcal{P}}(x_k^*; x_{obs}^0) \quad (3.42)$$

$$= \sum_{k=1}^K \mathcal{C}_{soc}^{\mathcal{P}}(x_k^*; x_{obs}^0) \cdot \frac{M_k}{M} \quad (3.43)$$

En supposant maintenant que le nombre d'utilisateurs devient tellement important que l'ensemble des valeurs admissibles par les valeurs cibles optimales devient continu, alors le terme $\frac{M_k}{M}$ tend vers une densité de probabilité, que nous noterons g (densité de probabilité des valeurs cibles optimales au sein des utilisateurs). Le coût effectif moyen se calcule alors par le biais d'une espérance : nous le renommons dans ce cas « *coût effectif sociétal* », que nous notons $E[\mathbf{Ceff}]$. L'équation 3.43 devient donc par passage à la limite :

$$E[\mathbf{Ceff}] = \int_{-\infty}^{+\infty} \mathcal{C}_{soc}^{\mathcal{P}}(x^*; x_{obs}^0) \cdot g(x^*) \cdot dx^* \quad (3.44)$$

Ce coût effectif sociétal constitue le score d'utilité que nous proposons d'utiliser. Dans la pratique, pour le calculer, il y a deux manières de procéder :

- ① Tous les utilisateurs ont été clairement identifiés et chacun d'eux a été capable de déterminer sa propre fonction de coût. Il est alors possible de calculer exactement la fonction de coût sociétale ainsi que la densité de probabilité g .
- ② Il n'est pas possible de définir de manière exhaustive toutes les fonctions de coût utilisées par l'ensemble des utilisateurs, soit parce que ceux-ci sont trop nombreux et dispersés pour être recensés, soit parce qu'ils sont incapables d'établir leurs fonctions de coût. On définit alors la fonction de coût sociétale directement à partir de l'observation de l'ensemble des utilisateurs ou de l'idée que l'on se fait de leurs besoins.

On ne connaît alors pas exactement la fonction g et on doit l'inférer elle aussi à partir de l'idée que l'on se fait des utilisateurs. Lorsque ceux-ci sont suffisamment nombreux et variés, il

peut être intéressant de considérer $g = f$, c'est à dire que les valeurs cibles optimales $x_{opt,j}^*$ des utilisateurs recouvrent l'ensemble des valeurs proposées par f , et ceci avec la même distribution. Cela revient à considérer que beaucoup d'utilisateurs ont intérêt à se préparer pour $\text{mode}(f)$, de moins en moins pour des valeurs de plus en plus éloignées de $\text{mode}(f)$ et seuls quelques rares utilisateurs se préparent pour les valeurs extrêmes.

③ Décomposition du coût effectif sociétal

Il est intéressant de considérer le coût effectif sociétal (eq. 3.44) pour différentes valeurs possibles de la réalisation : x_{obs}^0 varie alors selon la variable muette x^0 . On définit ainsi la fonction *coût sociétal attendu* dépendante de x^0 :

$$\mathbf{Asoc}(x^0) = \int_{-\infty}^{+\infty} \mathcal{C}_{soc}^{\mathcal{P}}(x^*; x^0) \cdot g(x^*) \cdot dx^* \quad (3.45)$$

Ce coût sociétal attendu correspond au coût moyen que subirait l'ensemble des utilisateurs si la valeur observée était x^0 . Nous avons bien sûr $E[\mathbf{Ceff}] = \mathbf{Asoc}(x_{obs}^0)$.

La fonction de coût sociétal respectant les contraintes d'une fonction de coût, \mathbf{Asoc} admet au moins un minimum, réalisé en $x^0 = x_{min}^0$:

$$\min \mathbf{Asoc} = \mathbf{Asoc}(x_{min}^0) \quad (3.46)$$

x_{min}^0 correspond à la valeur qui, si elle se réalise effectivement, permet à l'ensemble des utilisateurs de subir un coût global minimum. Nous l'appellerons *réalisation souhaitée*.

On peut donc définir l'écart $\Delta \mathbf{Asoc}$ entre ce minimum (quand $x^0 = x_{min}^0$) et le coût effectif sociétal (quand $x^0 = x_{obs}^0$), comme illustré par la figure 3.12. Nous avons ainsi la décomposition suivante :

$$E[\mathbf{Ceff}] = \min \mathbf{Asoc} + \Delta \mathbf{Asoc} \quad (3.47)$$

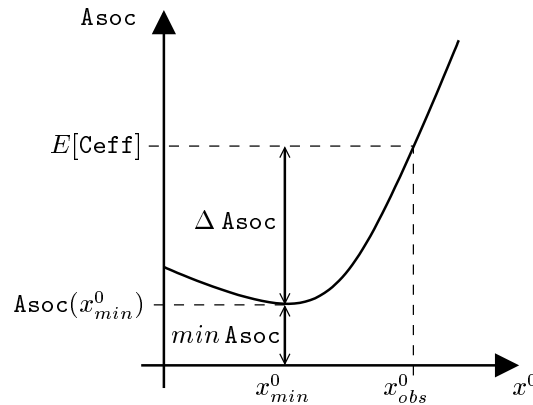


FIG. 3.12: Illustration de la décomposition du coût effectif sociétal.

Comme nous nous proposons de le montrer, $\Delta \mathbf{Asoc}$ est une mesure de la *justesse* de la prévision alors que $\min \mathbf{Asoc}$ est une mesure de sa *finesse* :

- De par sa définition, $\Delta \mathbf{Asoc}$ correspond à une mesure de l'écart entre x_{min}^0 et x_{obs}^0 .
Or, x_{min}^0 est la valeur qui permet d'obtenir le plus faible coût effectif sociétal : c'est donc ce que l'on souhaite voir prendre comme valeur par l'événement au vu de la prévision. $\Delta \mathbf{Asoc}$ est ainsi une mesure de l'écart entre ce qui est effectivement observé (x_{obs}^0) et ce que, compte-tenu de la prévision, on souhaiterait qu'il arrive (x_{min}^0) : c'est la justesse.
- Soit f_1 une prévision parfaitement fine, qui propose la seule valeur x_1^* :

$$\begin{cases} \forall x \neq x_1^* : f_1(x) = 0 \\ \int_{x_1^*}^{x_1^*} f_1(x).dx = 1 \end{cases} \quad (3.48)$$

Devant une telle prévision, nos utilisateurs n'ont pas le choix : ils doivent tous se préparer pour x_1^* , ce qui signifie que $g = f_1$. On a alors, en développant l'équation 3.45 : $\mathbf{Asoc}(x^0) = \mathcal{C}_{soc}^p(x_1^*; x^0)$. Etant donné que \mathcal{C}_{soc}^p obéit aux relations 3.32 et 3.31 qui contraignent les fonctions de coût, la fonction $\mathbf{Asoc}(x^0)$ est minimale et nulle pour $x^0 = x_1^*$.

Inversement, supposons que l'on ait une prévision f_1 telle que $\min \mathbf{Asoc} = 0$, minimum réalisé pour $x^0 = x_1^*$. Alors, d'après les relations 3.31 et 3.45, pour tout x^* tel que $g(x^*) \neq 0$, on a $\mathcal{C}_{soc}^p(x^*; x_1^*) = 0$. Or, en vertu de la relation 3.32, si $\mathcal{C}_{soc}^p(x^*; x_1^*) = 0$ alors $x^* = x_1^*$. Donc $g(x)$ est nulle pour tout x différent de x_1^* . Cela signifie que tous les utilisateurs se sont préparés pour la même valeur x_1^* . A condition que ceux-ci aient des fonctions de coût suffisamment diversifiées, cela implique que la prévision f_1 ne leur laissait pas le choix : elle est parfaitement fine et ne propose que la seule valeur x_1^* . Si les utilisateurs possèdent tous la même fonction de coût, nous nous retrouvons dans un cas similaire à celui d'un utilisateur unique, développé lors de l'étape ① : il nous est impossible de conclure sur la finesse de la prévision.

Il en résulte donc que, dès lors que les utilisateurs considérés ont des fonctions de coût suffisamment variées, $\min \mathbf{Asoc}$ est nul si et seulement si la prévision est parfaitement fine (et est strictement positive dans le cas contraire). Ainsi, $\min \mathbf{Asoc}$ est une mesure de la finesse de la prévision.

Le coût effectif sociétal est donc un score d'utilité qui permet de mesurer à la fois la justesse (par le biais de $\Delta \mathbf{Asoc}$) et la finesse des prévisions (par l'intermédiaire de $\min \mathbf{Asoc}$). La métrique associée à ces mesures est définie par le choix de la fonction de coût sociétal \mathcal{C}_{soc}^p .

④ Passage des scores individuels au score global

Nous n'avons considéré jusqu'à présent que le cas d'un unique couple prévision/observation. Notre score d'utilité $E[\mathbf{Ceff}]$ correspond donc à un Score Individuel.

Si nous nous plaçons maintenant dans le cas d'un échantillon de N couples prévision/observation, nous avons pour chaque couple i de l'échantillon une valeur du coût effectif sociétal que nous notons maintenant $E[Ceff]_i$. Pour obtenir un score d'utilité sur l'ensemble de l'échantillon, *i.e.*

un Score Global S , il suffit de composer ces Scores Individuels par une fonction h arbitrairement choisie (le plus souvent une moyenne arithmétique ou quadratique), comme exposé dans la partie 3.1.2 :

$$S = h((E[\mathbf{Ceff}]_i)_{i=1,N}) \quad (3.49)$$

En composant avec la même fonction h les finesesses individuelles $Finesse_i$ (resp. les justesses individuelles $Justesse_i$), on obtient la finesse globale $Finesse$ (resp. la justesse globale $Justesse$) sur l'ensemble de l'échantillon :

$$\begin{cases} Finesse &= h((Finesse_i)_{i=1,N}) \\ Justesse &= h((Justesse_i)_{i=1,N}) \end{cases} \quad (3.50)$$

3.3.5.2 Exemples choisis

Jusqu'ici, nous nous sommes placés dans le cadre général de fonctions de coût quelconques. Nous allons donc maintenant nous intéresser à quelques fonctions de coût usuelles et aux scores d'utilités associés.

Exemple 1

Dans un premier temps, nous allons considérer directement une fonction de coût sociétale, sans chercher à identifier de façon individuelle chaque utilisateur. Nous choisissons une fonction symétrique, c'est-à-dire qui pénalise une surestimation de la même façon qu'une sous-estimation. La fonction de coût symétrique la plus simple est bien sûr l'écart absolu entre la valeur cible et la réalisation :

$$\mathcal{C}_{soc}(x^*; x^0) = |x^* - x^0| \quad (3.51)$$

En l'absence de plus d'informations sur les utilisateurs, nous allons en outre identifier la densité de probabilité g des valeurs cibles optimales au sein des utilisateurs à celle donnée par la prévision (f).

Armés de cette caractérisation des utilisateurs (qui peut leur correspondre plus ou moins...), nous pouvons construire le Score Individuel d'utilité qui leur est associé (*i.e.* pour un unique couple prévision/observation). Le calcul du coût effectif sociétal s'effectue par la résolution de

l'équation 3.44 :

$$E[\mathbf{Ceff}] = \int_{-\infty}^{x_{obs}^0} (x_{obs}^0 - x^*) \cdot f(x^*) \cdot dx^* + \int_{x_{obs}^0}^{+\infty} (x^* - x_{obs}^0) \cdot f(x^*) \cdot dx^* \quad (3.52)$$

$$= x_{obs}^0 \cdot F(x_{obs}^0) - \int_{-\infty}^{x_{obs}^0} x^* \cdot f(x^*) \cdot dx^* + \int_{x_{obs}^0}^{+\infty} x^* \cdot f(x^*) \cdot dx^* - x_{obs}^0 \cdot (1 - F(x_{obs}^0)) \quad (3.53)$$

$$= x_{obs}^0 \cdot F(x_{obs}^0) - \left[x \cdot F(x) \right]_{-\infty}^{x_{obs}^0} + \int_{-\infty}^{x_{obs}^0} F(x^*) \cdot dx^* + \left[x \cdot (F(x) - 1) \right]_{x_{obs}^0}^{+\infty} - \int_{x_{obs}^0}^{+\infty} (F(x^*) - 1) \cdot dx^* - x_{obs}^0 \cdot (1 - F(x_{obs}^0)) \quad (3.54)$$

$$= \int_{-\infty}^{x_{obs}^0} F(x) \cdot dx + \int_{x_{obs}^0}^{+\infty} (1 - F(x)) \cdot dx \quad (3.55)$$

Le passage de l'équation 3.53 à l'équation 3.54 se fait par l'intermédiaire d'une intégration par parties. Encore une fois, nous coupons court aux questions concernant l'existence ou non de ces différentes intégrales en considérant que la densité de probabilité f est rigoureusement nulle hors d'un certain intervalle.

L'équation 3.55 peut finalement s'écrire sous la forme :

$$E[\mathbf{Ceff}] = \int_{-\infty}^{+\infty} \left| F(x) - H_{x_{obs}^0}(x) \right| \cdot dx \quad (3.56)$$

où $H_{x_{obs}^0}$ est la fonction de répartition d'une variable aléatoire qui prend toujours la valeur x_{obs}^0 :

$$H_{x_{obs}^0}(x) = \begin{cases} 0 & \text{si } x < x_{obs}^0 \\ 1 & \text{si } x \geq x_{obs}^0 \end{cases}$$

Le coût effectif sociétal s'interprète alors aisément graphiquement comme l'aire comprise entre $H_{x_{obs}^0}$ et la fonction de répartition F de la prévision (figure 3.13a).

Afin d'exprimer la décomposition de cette utilité en termes de finesse et de justesse, on recherche la valeur de la réalisation souhaitée x_{min}^0 en minimisant la fonction $\mathbf{Asoc}(x^0)$:

$$\frac{\partial \mathbf{Asoc}(x^0)}{\partial x^0} = 0 \quad \Leftrightarrow \quad F(x^0) - (1 - F(x^0)) = 0 \quad (3.57)$$

$$\Leftrightarrow \quad F(x^0) = \frac{1}{2} \quad (3.58)$$

$$\Leftrightarrow \quad x^0 = Q_{50\%}^F \quad (3.59)$$

où $Q_{50\%}^F$ est le quantile 50% de la prévision F (ce qui correspond à la médiane). Nous avons donc :

$$\text{Finesse} = \min \mathbf{Asoc} = \int_{-\infty}^{+\infty} \left| F(x) - H_{Q_{50\%}^F}(x) \right| \cdot dx \quad (3.60)$$

$$\text{Justesse} = \Delta \mathbf{Asoc} = \int_{x_{obs}^0}^{Q_{50\%}^F} (1 - 2 \cdot F(x)) \cdot dx \quad (3.61)$$

Finesse et Justesse s'interprètent donc aussi aisément de façon graphique (figure 3.13b et c)

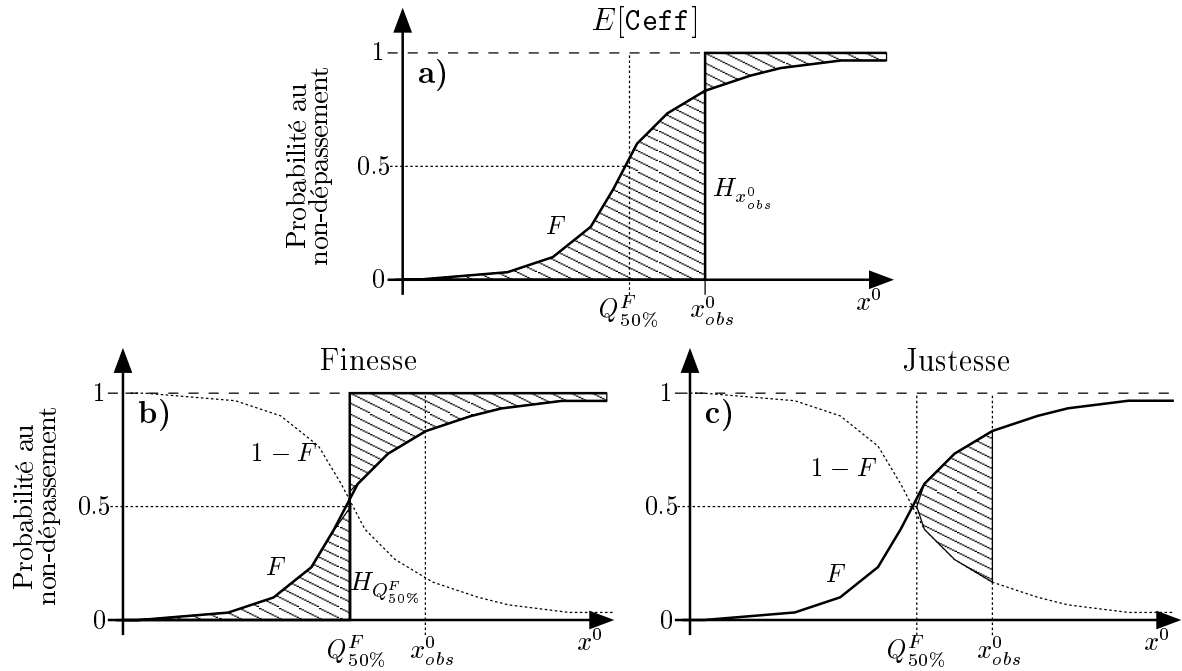


FIG. 3.13: Interprétation graphique du score d'utilité et de sa décomposition en finesse et justesse. **a)** Le coût effectif sociétal correspond à l'aire comprise entre la fonction de répartition F donnée par la prévision et la fonction $H_{x_{obs}^0}$, **b)** la finesse à l'aire comprise entre F et la fonction $H_{Q_{50\%}^F}$ et **c)** la justesse à l'aire comprise entre les fonctions de répartition F et $1 - F$, de $Q_{50\%}^F$ à x_{obs}^0 .

Pour obtenir un Score Global, on peut par exemple composer le Score Individuel obtenu par une moyenne arithmétique :

$$S = \frac{1}{N} \sum_{i=1}^N \left(\int_{-\infty}^{+\infty} |F_i(x) - H_{x_{obs,i}^0}(x)| . dx \right) \quad (3.62)$$

où l'indice i est l'index des N éléments de l'échantillon d'événements sur lequel est calculé S . Ce résultat est très proche du *CRPS* (équation 3.18). Il présente néanmoins des caractéristiques fort différentes car il est biaisé. Le lecteur pourra en effet vérifier que s'il considère une loi de probabilité prévue unique sur un échantillon quelconque, la prévision déterministe contituée par la moyenne des observations sur l'échantillon obtient un meilleur score que la prévision de la climatologie de l'échantillon.

Exemple 2

Nous nous intéressons maintenant à un ensemble d'utilisateurs pour lesquels la fonction de coût de chacun d'entre eux est connue. Nous allons ainsi construire la fonction de coût sociétale résultante afin de pouvoir exprimer le coût effectif sociétal.

Après interview d'un panel représentatif, il s'avère que notre ensemble d'utilisateurs peut se définir ainsi :

- Chaque utilisateur j possède une fonction de coût de la forme suivante :

$$\mathcal{C}_j(x^*; x^0) = \begin{cases} (1 + k_j) \cdot (x^0 - x^*) & \text{si } x^0 > x^* \\ (1 - k_j) \cdot (x^* - x^0) & \text{si } x^0 \leq x^* \end{cases} \quad (3.63)$$

Le paramètre k_j est propre à chaque utilisateur. Il est compris entre -1 et 1, chacune de ces valeurs correspondant à un cadre décisionnel limite (figure 3.14) :

- $k_j = 1$: une surestimation n'implique aucun coût, alors qu'une sous-estimation a un coût non nul. L'utilisateur se prépare donc toujours pour la plus grande valeur possible.
- $k_j = -1$: à l'inverse du cas précédent, une sous-estimation est sans conséquence sur le coût alors qu'une surestimation coûte cher. L'utilisateur se prépare donc toujours pour la plus petite valeur possible.

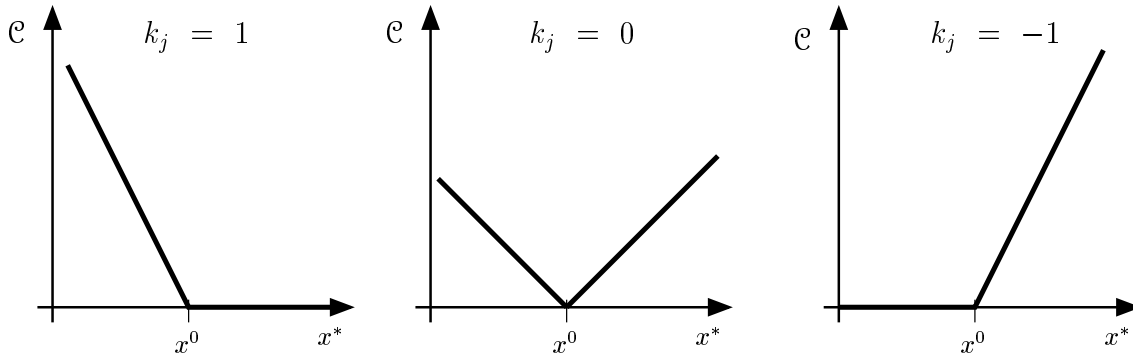


FIG. 3.14: Représentation de la fonction de coût pour x^0 fixée, en fonction de x^* .

- la distribution de nos utilisateurs, entre ceux qui sont beaucoup ou peu pénalisés par une surestimation ou une sous-estimation, est telle que le paramètre k_j suit une loi de distribution uniforme entre -1 et 1.

Nous allons cette fois encore commencer par considérer tout d'abord un unique couple prévision/observation.

La résolution de l'équation 3.33 permet de calculer l'espérance attendue du coût de chaque utilisateur (par un raisonnement similaire à celui des équations 3.52 à 3.55) :

$$E_j^{\mathcal{C}, \mathcal{P}}(x^*) = (1 - k_j) \int_{-\infty}^{x^*} F(x^0) \cdot dx^0 + (1 + k_j) \int_{x^*}^{+\infty} (1 - F(x^0)) \cdot dx^0 \quad (3.64)$$

On recherche alors le minimum de cette fonction afin d'exprimer la valeur cible optimale de chaque utilisateur :

$$\frac{\partial E_j^{\mathcal{C}, \mathcal{P}}(x^*)}{\partial x^*} = 0 \quad \Leftrightarrow \quad F(x^*) = \frac{1 + k_j}{2} \quad (3.65)$$

d'où :

$$x_{opt,j}^* = F^{-1} \left(\frac{1 + k_j}{2} \right) \quad (3.66)$$

Afin de définir la fonction de coût sociétale, considérons une valeur fixée de x_1^* de $x_{opt,j}^*$. En vertu de l'équation 3.66, tous les utilisateurs qui acceptent x_1^* comme valeur cible optimale, sont ceux qui possèdent un paramètre k tel que :

$$k = 2F(x_1^*) - 1 \quad (3.67)$$

Ce sont donc les utilisateurs qui possèdent la fonction de coût suivante (obtenue en reportant 3.67 dans 3.63) :

$$\mathcal{C}(x_1^*; x^0) = \begin{cases} 2.F(x_1^*). (x^0 - x_1^*) & \text{si } x^0 > x_1^* \\ 2.(1 - F(x_1^*)). (x_1^* - x^0) & \text{si } x^0 \leq x_1^* \end{cases} \quad (3.68)$$

En laissant varier x_1^* , nous obtenons l'expression de la fonction de coût sociétale :

$$\mathcal{C}_{soc}^{\mathcal{P}}(x^*; x^0) = \begin{cases} 2.F(x^*). (x^0 - x^*) & \text{si } x^0 > x^* \\ 2.(1 - F(x^*)). (x^* - x^0) & \text{si } x^0 \leq x^* \end{cases} \quad (3.69)$$

Pour calculer le score d'utilité, il ne reste plus qu'à calculer le coût effectif sociétal par l'intermédiaire de l'équation 3.44. Le paramètre k_j étant réparti de façon uniforme entre -1 et 1 au sein de nos utilisateurs, la relation 3.65 implique que $F(x_{opt,j}^*)$ suit une loi de probabilité uniforme entre 0 et 1. Il en résulte donc que la variable aléatoire x_{opt}^* est régie par la densité de probabilité f . Nous avons donc $g = f$, d'où :

$$E[\mathbf{Ceff}] = \int_{-\infty}^{+\infty} \mathcal{C}_{soc}^{\mathcal{P}}(x^*; x_{obs}^0). f(x^*). dx^* \quad (3.70)$$

$$= \int_{-\infty}^{x_{obs}^0} F^2(x). dx + \int_{x_{obs}^0}^{+\infty} (1 - F(x))^2. dx \quad (3.71)$$

$$= \int_{-\infty}^{+\infty} \left(F(x) - H_{x_{obs}^0}(x) \right)^2. dx \quad (3.72)$$

où $H_{x_{obs}^0}$ est défini comme dans l'exemple précédent.

Dans le cas où l'on considère un Score Global fondé sur la composition selon une moyenne arithmétique des Scores Individuels ainsi définis, on retrouve l'expression du CRPS (cf. 3.3.3.3). Ce score est donc bien une mesure de l'utilité des prévisions, même s'il est couramment utilisé comme score de qualité.

En recherchant la valeur de la réalisation souhaitée, via la minimisation de la fonction $\mathbf{Asoc}(x^0)$, nous sommes donc en mesure de proposer une décomposition du CRPS en terme de Finesse et de Justesse :

$$\frac{\partial \mathbf{Asoc}(x^0)}{\partial x^0} = 0 \quad \Leftrightarrow \quad x^0 = Q_{50\%}^F \quad (3.73)$$

où, comme dans l'exemple précédent, $Q_{50\%}^F$ est le quantile 50% de la prévision F .

Ceci nous conduit donc aux expressions suivantes :

$$\text{Finesse} = \min \mathbf{A} \mathbf{s} \mathbf{o} \mathbf{c} = \int_{-\infty}^{+\infty} \left(F(x) - H_{Q_{50\%}^F}(x) \right)^2 dx \quad (3.74)$$

$$\text{Justesse} = \Delta \mathbf{A} \mathbf{s} \mathbf{o} \mathbf{c} = \int_{x_{obs}^0}^{Q_{50\%}^F} (1 - 2.F(x)).dx \quad (3.75)$$

Là encore, finesse et justesse s'interprètent graphiquement (figure 3.15).

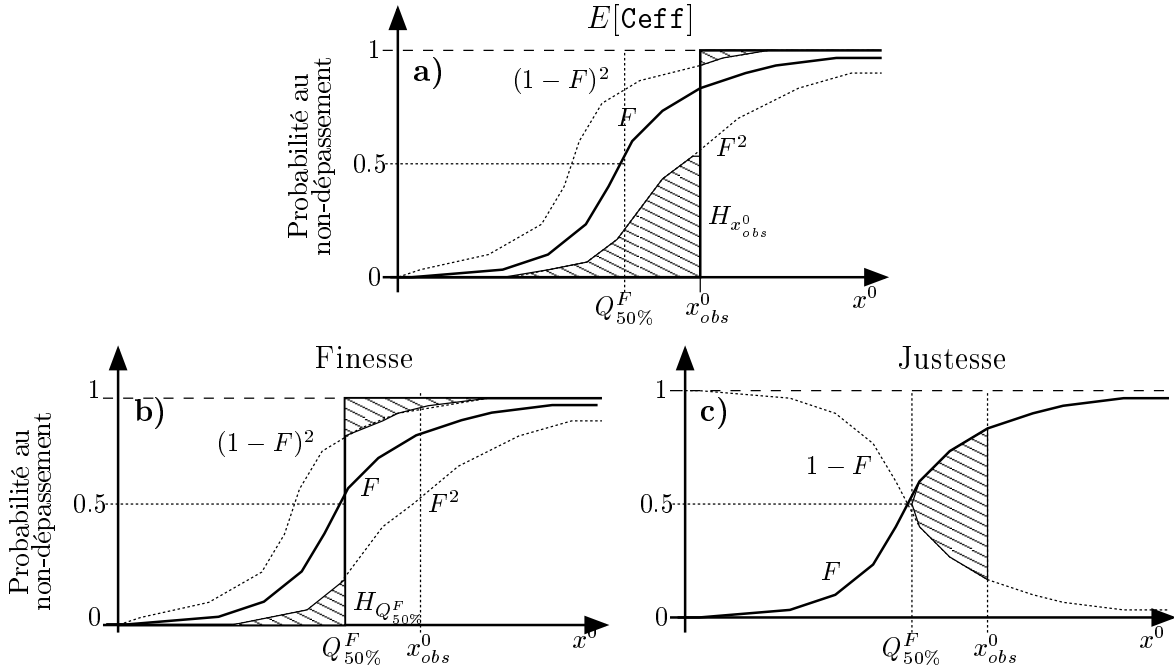


FIG. 3.15: Interprétation graphique : (a) du score d'utilité CRPS, et de sa décomposition en : (b) finesse et (c) justesse.

On remarque que la mesure de la justesse est identique à celle de l'exemple précédent. Seule la mesure de la finesse diffère. La figure 3.16 présente quelques exemples de l'évolution des coûts effectifs sociétaux ainsi obtenus en fonction de x_{obs}^0 , pour différentes prévisions de même justesse. Ces courbes permettent d'illustrer le fonctionnement du *CRPS* : plus la réalisation effectivement observée (x_{obs}^0) est éloignée de la réalisation souhaitée ($Q_{50\%}^F$), plus l'excès de finesse (*i.e.* la surestimation de ses capacités prédictives) est pénalisé.

3.4 Choix arbitrairement retenus

Dans le cadre du travail présenté ici, nous sommes amenés à utiliser des scores de vérification dans deux contextes distincts d'applications :

- d'une part, en calibration, lors de l'optimisation de la méthode de prévision et des tests de nouveaux algorithmes, afin de les comparer et de choisir le meilleur,

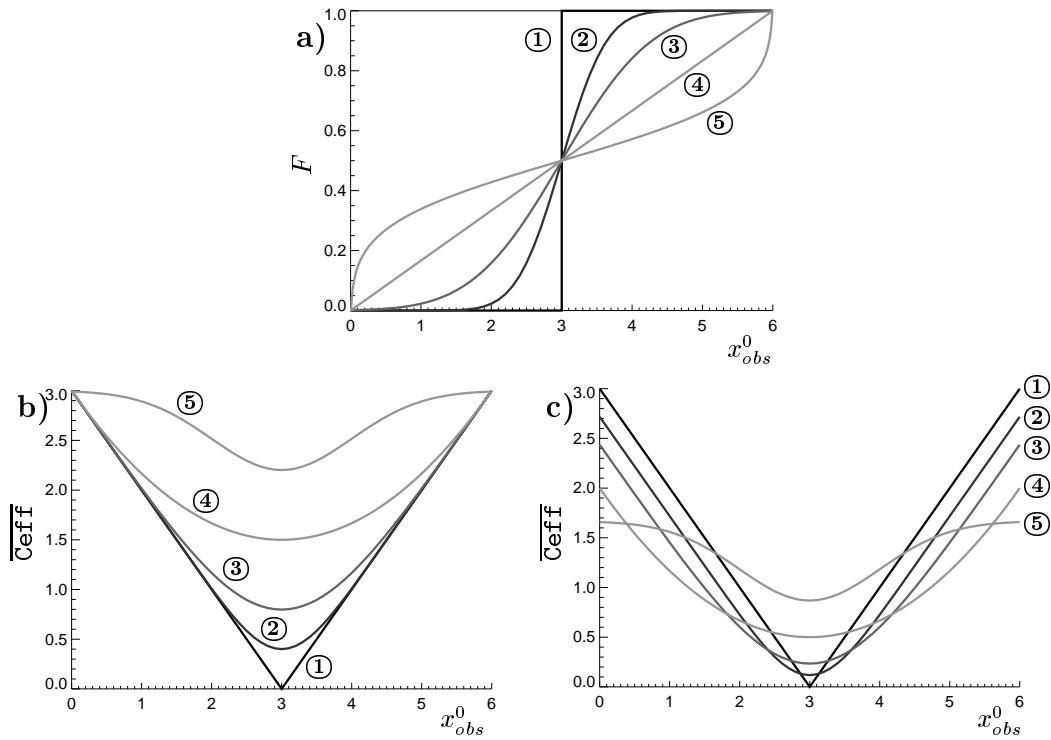


FIG. 3.16: Evolution des coûts effectifs sociétaux en fonction de x_{obs}^0 pour différentes prévisions (présentées dans le panneau (a)) : (b) issus de l'exemple 1, (c) issus de l'exemple 2

– et d'autre part, en validation, lors du diagnostic des performances des algorithmes finalement retenus.

Ces deux types d'utilisation conduisent à des choix différents en matière de scores de vérification.

3.4.1 L'optimisation des algorithmes

La phase d'optimisation implique le test d'un très grand nombre de configurations différentes d'algorithmes, ce qui génère autant de séries de prévisions. Il n'est donc pas pensable dans un tel contexte d'utiliser des critères de vérification qui aient besoin d'interprétations humaines. Le critère utilisé doit donc être unique : à chaque série de prévisions correspond un et un seul score de qualité/utilité. La série qui présente le meilleur score est alors la plus performante et l'algorithme qui l'a produite sera retenu comme le meilleur. En outre, ce critère ne doit pas mettre en avant un seul aspect de la prévision mais doit être un intégrateur qui prend en compte tous les attributs désirables de la prévision.

Lors de ses travaux en 1997, Guilbaud a utilisé le *RPS*. Nous nous proposons donc, dans la continuité de ce travail, de considérer la version continue du *RPS* : le *CRPS* (cf. 3.3.3.3). De cette manière, nous sommes dispensés de l'étape toujours délicate de la définition de classes de pluie. Le CRPS est un score non-biaisé et les développements théoriques précédents (cf. 3.3.5.1) ont montré qu'il mesure à la fois la Finesse et la Justesse des prévisions.

De plus, comme nous venons de le montrer, ce score mesure le coût effectif sociétal d'un ensemble d'utilisateurs j ayant chacun la fonction de coût suivante :

$$\mathcal{C}_j(x^*; x^0) = \begin{cases} (1 + k_j) \cdot (x^0 - x^*) & \text{si } x^0 > x^* \\ (1 - k_j) \cdot (x^* - x^0) & \text{si } x^0 \leq x^* \end{cases} \quad (3.76)$$

le paramètre k_j étant réparti uniformément au sein de l'ensemble des utilisateurs. Le *CRPS* mesure donc l'utilité des prévisions pour ce panel disparate d'utilisateurs. Or, le travail réalisé ici n'est pas destiné à une utilisation particulière mais doit pouvoir convenir à un ensemble d'utilisateurs le plus large possible, tel que celui sur lequel repose le *CRPS*. C'est pourquoi l'utilisation de ce score nous paraît appropriée à nos objectifs.

Afin de donner plus de signification aux valeurs numériques obtenues, nous les exprimerons toujours, sauf indication contraire, en terme de compétence par rapport à la climatologie (cf. 3.1.4).

3.4.2 Le diagnostic des performances finales

Une fois la phase d'optimisation achevée, nous ne disposerons plus que de deux ou trois algorithmes. Il nous sera alors loisible d'utiliser des scores de vérification qui ne jugent qu'un aspect des prévisions obtenues en vue d'établir un bilan détaillé de leurs performances.

Nous mettrons ainsi en œuvre les scores issus de l'usage de tableaux de contingence : ceci permettra de mettre en évidence les spécificités des prévisions obtenues en fonction des *probabilités de coupure* et des *valeurs seuils* considérées.

Deuxième partie

Optimisation de la méthode

Introduction à la PARTIE II

L'objet de cette partie est de définir quelle est la méthode d'adaptation par analogie la plus performante, en matière de prévision de précipitations journalières.

Pour ce faire, nous allons tester différentes façons de définir une analogie, en considérant successivement divers *algorithmes*. Sur une même série d'événements, chacun d'eux va produire sa propre série de prévisions. La comparaison de l'adéquation entre chacune de ces prévisions et les observations correspondantes, effectuée par l'intermédiaire de la compétence du *CRPS* sur la climatologie (cf. 3.4), permettra de définir l'algorithme le plus utile, donc le plus performant. Rappelons ici que nous ne travaillons pas directement avec les précipitations brutes, mais avec les valeurs de précipitations transformées selon l'équation 2.2.

Il est important de bien noter que la prépondérance d'un algorithme plutôt qu'un autre est directement liée au prédictand choisi. Un tel algorithme est spécifique à la variable cible. Nous nous intéressons ici aux précipitations cumulées sur 24 heures, mais il est évident que si l'on s'intéresse à d'autres prédictands (précipitations cumulées sur 12 ou 6 heures, durées d'insolation, rafales de vents...), même si l'approche globale reste applicable, l'algorithme le plus performant sera différent.

Etant donné que nous recherchons le meilleur algorithme d'adaptation, indépendamment de tout modèle NWP, nous nous plaçons ici en prévision parfaite. Les données des situations météorologiques pour lesquelles nous rechercherons des analogues (situations cibles) seront donc issues des réanalyses NCEP/NCAR de l'archive LTRE, à l'instar des données des journées potentiellement analogues (situations candidates).

En outre, afin d'être en mesure de vérifier dans la partie III que les résultats que nous allons obtenir ne sont pas le simple fruit du hasard, liés au choix des situations cibles, nous avons scindé notre archive en deux parties :

- **La période de calibration**, qui sert à optimiser les algorithmes d'adaptation par analogie :
 - situations cibles : de 1953 à 1996 (soit 16070 événements),
 - situations candidates : de 1953 à 1996.
- **La période de validation**, qui permet de vérifier la pertinence de l'optimisation en s'assurant de sa robustesse sur un jeu de données indépendant :
 - situations cibles : de 1997 à 2001 (soit 1825 événements),
 - situations candidates : de 1953 à 1996.

Tout au long de cette partie, nous allons en quelque sorte « *rejouer* » les réanalyses. Evidemment, étant donné que les situations cibles couvrent la même période que les situations candidates, la meilleure analogue à n'importe quelle situation cible sera elle-même ! Pour remédier à ce problème, nous nous interdisons de rechercher des analogues dans la même année que la situation cible. Ainsi, bien que l'archive des situations candidates couvre 44 années, seules 43 de ces années sont effectivement utilisées pour chaque prévision. Par ailleurs, sauf précision contraire, nous adoptons le principe de la fenêtre glissante de plus ou moins 2 mois utilisée par S. Guilbaud (1997) pour prendre en compte les effets saisonniers. Cela nous autorise donc à passer en revue en moyenne $43 \times 2 \times 60 = 5160$ situations candidates pour chacune des 16070 situations cibles, comme illustré ci-dessous.

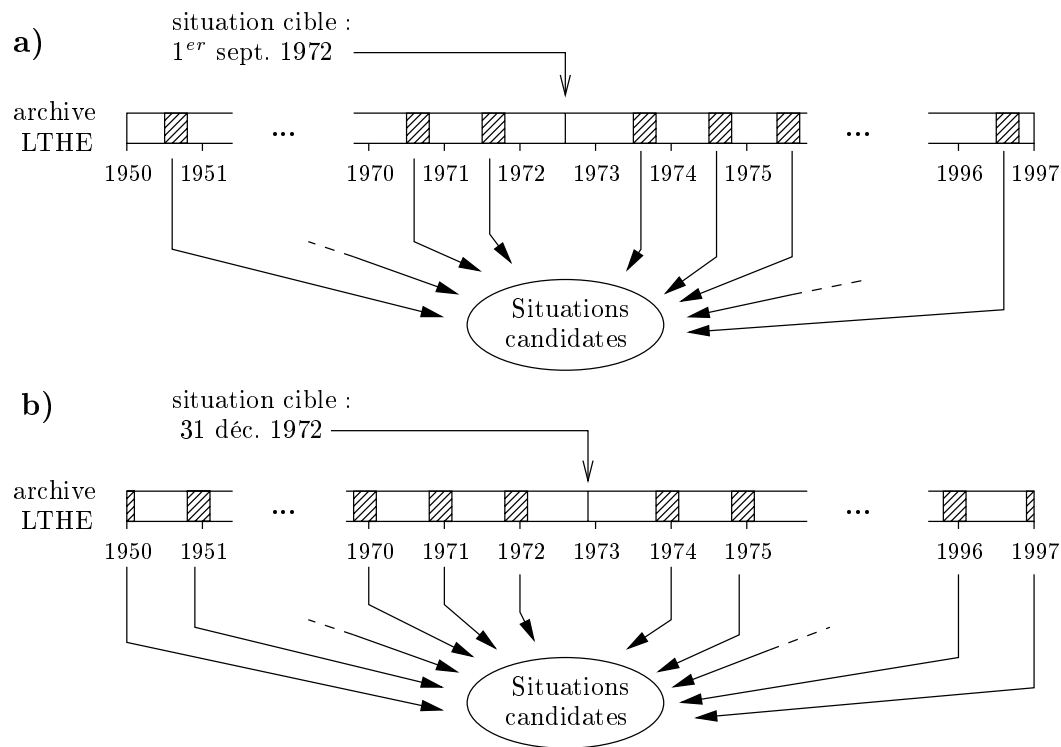


Schéma représentant la localisation dans l'archive LTHE des situations candidates pour une situation cible correspondant à **a)** 1^{er} septembre 1972, et **b)** 31 décembre 1972.

4

Les géopotentiels comme variable d'analogie principale

« *Glao a zeu diwar greisteiz, glao e pad an deiz.
Glao diouz ar viz, glao e pad ar miz.* »

« *Pluie qui vient du sud, pluie toute la journée.
Pluie du nord-est, pluie tout le mois.* »

Dicton populaire Breton

Ce dicton breton est une sorte de prévision par analogie. La mémoire populaire a ainsi retenu une liaison entre la direction du vent (la variable d'analogie), et la durée de la pluie (le prédicand). Malheureusement, la mémoire humaine est faillible : elle est trop sélective pour que l'on puisse construire dessus des relations météorologiques fiables.

Néanmoins, même si la justesse de ce dicton est sujette à caution, il demeure l'idée d'une liaison entre les précipitations et certains éléments synoptiques de l'atmosphère. C'est cette liaison que nous allons étudier ici de façon systématique, en recherchant quelle est la variable d'analogie la plus performante pour la prévision des précipitations journalières sur les groupements de l'archive pluviométrique EDF.

Nous allons tout d'abord définir plusieurs critères d'analogie que nous appliquerons à chacune des variables présentes au sein de l'archive météorologique LTHE. Nous retiendrons ainsi le couple variable d'analogie/critère d'analogie qui définit l'analogie la plus pertinente. Dans un second temps, nous étudierons en détail l'influence du choix de l'heure d'observation ainsi que du niveau géopotentiel de mesure de cette variable, dans le contexte d'une analogie portant sur un unique champ météorologique. L'étude de la composition de plusieurs champs de même nature conclura finalement ce chapitre.

Durant tout ce chapitre, les fenêtres d'analogie utilisées seront imposées de façon arbitraire. L'étude détaillée de leur influence et de leur optimisation fera l'objet du chapitre suivant. On perçoit ici l'aspect combinatoire de notre problématique : variables et fenêtres d'analogie ne sont pas indépendantes. Nous présentons nos résultats en recherchant d'abord la meilleure variable

sur une fenêtre imposée, puis la meilleure fenêtre pour la variable retenue. Cependant, même en ré-itérant le processus, est-on sûr d'avoir trouvé la combinaison optimale ?

4.1 Exploration préliminaire

Toutes les variables disponibles dans l'archive météorologique LTRE ne sont pas forcément pertinentes vis-à-vis de notre variable cible spécifique : les précipitations journalières. Certaines sont sûrement plus informatives que d'autres, car associées plus directement aux processus générateurs de pluie. Nous allons donc passer en revue toutes les variables de l'archive, sans a priori, afin d'identifier les plus performantes.

4.1.1 Les critères d'analogies

Pour mettre en œuvre une recherche d'analogies, il est nécessaire de disposer d'un critère d'analogie. Celui-ci permet de quantifier le niveau de ressemblance entre deux situations, par rapport à la variable d'analogie considérée.

Il n'existe toutefois pas de critères d'analogie universels. En effet, une analogie peut se définir par l'intermédiaire de deux paramètres plus ou moins indépendants, que le critère d'analogie doit résumer en une seule valeur :

- La **distance moyenne** entre les deux champs, qui mesure leur écart moyen point à point : c'est un critère de *position*.
- Et leur **co-fluctuation**, qui évalue si ceux-ci varient de la même façon dans l'espace : c'est pour sa part un critère de *forme*.

En schématisant trois champs de hauteurs géopotentielles, la figure 4.1 permet d'illustrer la différence entre ces attributs. Considérons le champ **A** comme celui de la journée cible et les champs **B** et **C** comme ceux des journées candidates :

- **B** a exactement la même forme que **A**, mais se trouve à un niveau moyen inférieur (journée plus froide),
- **C** a une forme opposée à **A**, mais se trouve à la même altitude moyenne.

La meilleure analogue à la situation **A** est donc **C**, si l'on considère la distance moyenne comme critère d'analogie, alors que ce serait plutôt **B** au sens de la co-fluctuation. Ainsi, selon que l'on donne plus de poids à l'un ou l'autre de ces attributs, on sera amené à utiliser un critère d'analogie plutôt qu'un autre.

Les critères que nous citons ci-après s'appliquent tous à des champs définis aux nœuds d'une grille régulière. Nous supposons dans la suite que la fenêtre d'analogie possède I points en longitude et J en latitude. Nous noterons $V_{i,j}^{cible}$ (resp. $V_{i,j}^{candidate}$) la valeur de la variable d'analogie de la journée cible (resp. candidate) au point de grille de coordonnées (i, j) .

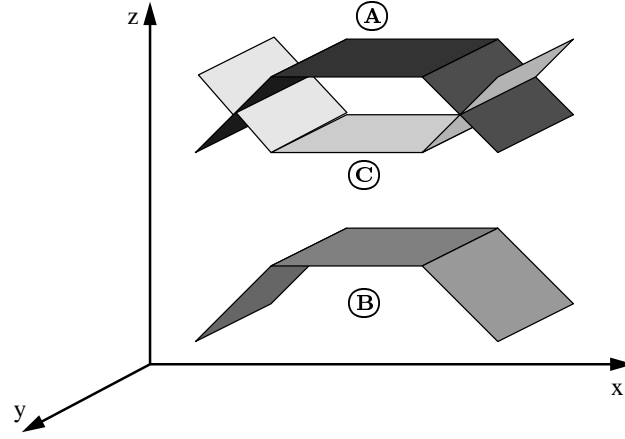


FIG. 4.1: Schéma d'un champ de hauteurs géopotentiellelles pour trois situations différentes.

4.1.1.1 Critères liés à la distance moyenne

Nous retrouvons ici les scores MAE et $RMSE$ définis précédemment (cf. 3.2.2.2) mais appliqués cette fois-ci à la comparaison de champs à deux dimensions :

$$MAE = \frac{1}{I.J} \sum_{i=1}^I \sum_{j=1}^J |V_{i,j}^{cible} - V_{i,j}^{candidate}| \quad (4.1)$$

$$RMSE = \sqrt{\frac{1}{I.J} \sum_{i=1}^I \sum_{j=1}^J (V_{i,j}^{cible} - V_{i,j}^{candidate})^2} \quad (4.2)$$

4.1.1.2 Critères liés à la co-fluctuation

Le critère de mesure de la co-fluctuation le plus évident est le coefficient de corrélation r que nous avons déjà abordé précédemment (cf. 3.2.2.3). Nous rappelons ici son expression dans le cas de la comparaison de deux situations météorologiques (cible et candidate) sur le même champ V :

$$r = \frac{\frac{1}{I.J-1} \cdot \sum_{i=1}^I \sum_{j=1}^J (V_{i,j}^{cible} - m_{V^{cible}})(V_{i,j}^{candidate} - m_{V^{candidate}})}{s_{V^{cible}} \cdot s_{V^{candidate}}} \quad (4.3)$$

où :

- $m_{V^{cible}}$ est la moyenne arithmétique des $V_{i,j}^{cible}$ sur la fenêtre d'analogie,
- $m_{V^{candidate}}$ est la moyenne arithmétique des $V_{i,j}^{candidate}$,
- $s_{V^{cible}}$ est l'écart-type des $V_{i,j}^{cible}$,
- $s_{V^{candidate}}$ est l'écart-type des $V_{i,j}^{candidate}$.

Bien que couramment utilisée, la corrélation possède deux propriétés qui peuvent devenir des inconvénients dans le cas présent :

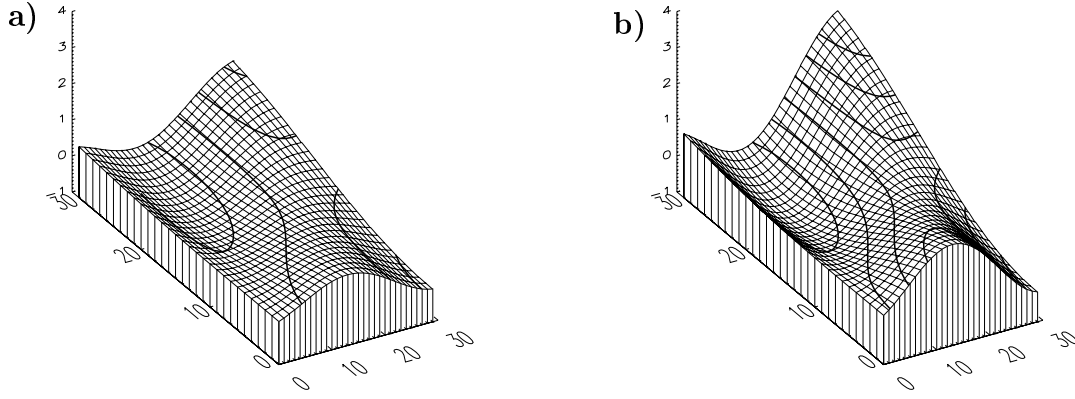


FIG. 4.2: Exemple de deux surfaces parfaitement corrélées ($r = 1$) mais présentant un facteur d'échelle différent de 1.

- La corrélation est insensible à un changement d'échelle sur l'une des séries. Ainsi, si l'un des champs se déduit de l'autre par un simple facteur multiplicatif appliqué identiquement à tous les points de grille, on aura $r = 1$ alors que l'un des champs est manifestement « étiré » par rapport à l'autre (figure 4.2).
- La corrélation perd par ailleurs une partie de l'information contenue dans la disposition spatiale des points de mesure : les différents points constitutifs du champ sont considérés indépendamment les uns des autres. A titre d'exemple, considérons la figure 4.3 présentant une situation cible (a) et deux situations candidates (b) et (c). Le champ (b) ne diffère du champ (c) que par l'intervention de certains de ses points. Celle-ci est effectuée de telle façon que les couples $(V_{i,j}^{cible}; V_{i,j}^{candidate})$ soient identiques pour les deux situations candidates. La corrélation ne tenant pas compte de la répartition spatiale des couples de valeur cible et candidate, le coefficient de corrélation avec (a) est ainsi le même pour (b) et (c), à savoir $r = 0.905$. La corrélation n'est donc pas apte à juger complètement de la cohérence spatiale de la co-fluctuation.

D'autres critères d'analogies ont donc également été proposés, et notamment le critère S1 (Teweless et Wobus 1954). Ce critère a été introduit afin de vérifier la qualité des prévisions de champs de hauteurs géopotentiels. Il s'attache à considérer les gradients des champs que l'on veut comparer, et non pas leurs seules valeurs ponctuelles comme le fait la corrélation point à point. Il s'exprime de la manière suivante :

$$S1 = 100. \frac{\sum_{i=1}^{I-1} \sum_{j=1}^J \left| \Delta_{i,j}^{i,cible} - \Delta_{i,j}^{i,candidate} \right| + \sum_{i=1}^I \sum_{j=1}^{J-1} \left| \Delta_{i,j}^{j,cible} - \Delta_{i,j}^{j,candidate} \right|}{\sum_{i=1}^{I-1} \sum_{j=1}^J \max \left(\left| \Delta_{i,j}^{i,cible} \right|; \left| \Delta_{i,j}^{i,candidate} \right| \right) + \sum_{i=1}^I \sum_{j=1}^{J-1} \max \left(\left| \Delta_{i,j}^{j,cible} \right|; \left| \Delta_{i,j}^{j,candidate} \right| \right)} \quad (4.4)$$

où :

- $\Delta_{i,j}^{i,cible}$ est le gradient dans la direction est-ouest selon le sens positif à partir du point (i, j) du champ cible : $\Delta_{i,j}^{i,cible} = V_{i+1,j}^{cible} - V_{i,j}^{cible}$
- $\Delta_{i,j}^{j,cible}$ est le gradient dans la direction nord-sud selon le sens positif à partir du point (i, j) du champ cible : $\Delta_{i,j}^{j,cible} = V_{i,j+1}^{cible} - V_{i,j}^{cible}$

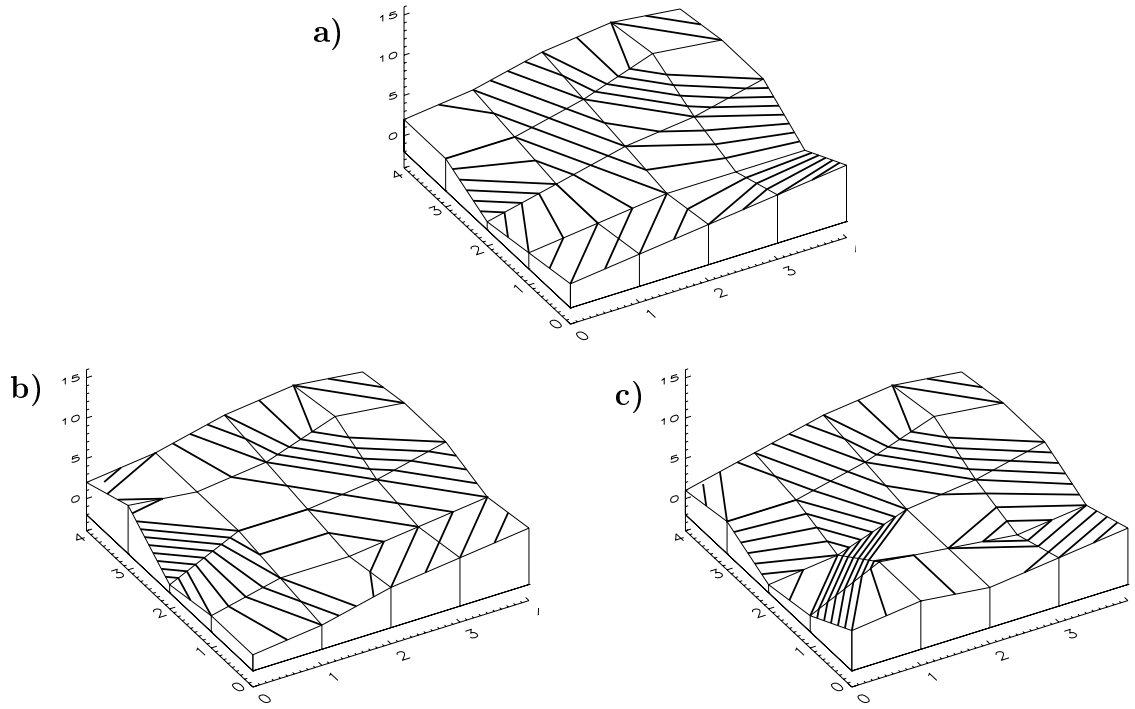


FIG. 4.3: Exemple de trois champs de hauteur que l'on souhaite comparer par corrélation. **a)** : situation cible; **b)** et **c)** : situations candidates.

Ce critère calcule donc la somme des écarts des gradients entre les situations cibles et les situations candidates (numérateur). Le rôle du dénominateur est simplement de normaliser cette somme d'écarts en prenant en compte, pour chacun d'eux, la valeur maximale des gradients considérés.

Le score $S1$ est sans dimension. Il varie de 0, pour deux champs identiques (figure 4.4a), à 200, pour des champs en totale opposition (figure 4.4c). Une valeur de 100 correspond à deux champs indépendants (figure 4.4b).

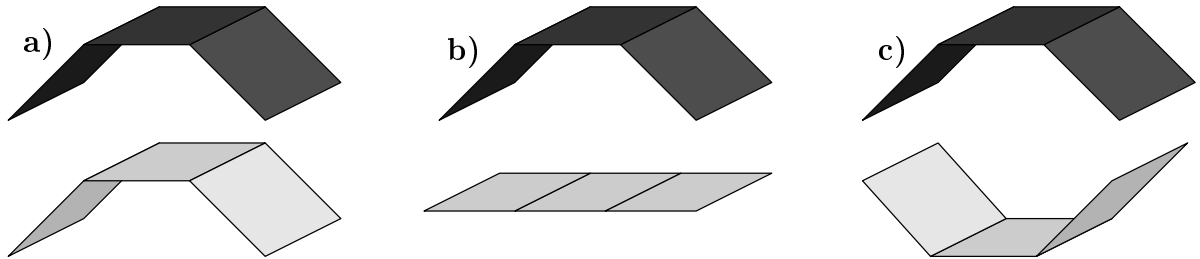


FIG. 4.4: Comparaison de champs à l'aide du score $S1$. **a)** : $S1 = 0$; **b)** : $S1 = 100$; **c)** : $S1 = 200$.

A l'instar de la corrélation, le score $S1$ est indépendant de la distance moyenne entre les champs cible et candidat. Par contre, il est sensible aux changements d'échelle et prend en compte la répartition spatiale des points de mesure les uns par rapport aux autres. Ainsi, dans le cas de l'exemple de la figure 4.3, nous avons $S1 = 47$ entre (a) et (b) et $S1 = 41$ entre (a)

et (c). La meilleure analogue à (a) est donc la situation (c), car, en quelque sorte, c'est celle qui « s'emboîte » le mieux avec (a).

4.1.2 Résultats préliminaires

En considérant successivement les critères d'analogie $RMSE$, r et $S1$, nous avons testé une à une les variables d'analogie présentes dans l'archive LTHE. Les situations cibles et candidates sont celles de la période de calibration, comme expliqué dans l'introduction à cette partie II.

Les performances obtenues sont présentées ci-après en ne considérant, pour chaque type de variable, que le niveau et l'heure d'observation qui donnent les meilleurs résultats. La figure 4.5a présente la performance moyenne sur l'ensemble des groupements EDF des 50 meilleures analogues, en considérant la fenêtre d'analogie utilisée par la méthode de référence (fenêtre de $19.1^\circ \times 11.3^\circ$, figure 1.9), tandis que la figure 4.5b présente les résultats obtenus en considérant une fenêtre de seulement $2.5^\circ \times 2.5^\circ$, centrée sur chaque groupement (chaque champ est alors défini par 4 points).

Sans surprise, la meilleure variable d'analogie est très nettement l'altitude du géopotentiel (HGT), et ce quel que soit le critère d'analogie considéré. Nous retrouvons les résultats obtenus par S. Guilbaud, en montrant que pour cette variable, le critère le plus pertinent est $S1$. Or, les gradients de géopotentiels déterminent le vent géostrophique¹, c'est à dire la circulation générale de l'atmosphère. Pour expliquer les précipitations, la meilleure analogie est donc bien une analogie sur la circulation synoptique, comme cela a été mis à profit par notre dicton breton.

L'analogie portant sur les intensités du vent zonal (UGR) ou méridien (VGR) est moins performante car elle ne prend en compte qu'un aspect de cette circulation : la connaissance de l'intensité du vent zonal n'est par exemple pas suffisante pour connaître la direction du vecteur vent. En outre, le vent réel (surtout dans les basses couches) subit l'influence de la topographie et est donc moins représentatif des processus à grande échelle. A l'inverse, les gradients de hauteurs géopotentielle définissent totalement les grandes lignes de la circulation atmosphérique.

Il convient de noter que quelle que soit la variable d'analogie considérée, l'utilisation de la corrélation comme critère d'analogie est systématiquement moins performante que celle du score $S1$ (sauf dans un seul cas, pour lequel la différence n'est pas significative). Par contre, le critère de distance $RMSE$ se montre très fréquemment plus intéressant que le score $S1$: ce dernier n'est réellement pertinent que pour la comparaison des champs de géopotentiels. Il semblerait que HGT soit la seule variable pour laquelle l'analogie sur la co-fluctuation est largement prépondérante sur celle concernant la distance moyenne.

Les variables définies suivant des niveaux isentropiques (UG2, VG2, TM2) se montrent toujours moins informatives que leurs homologues définies selon des niveaux géopotentiels (UGR, VGR, TMP). Ce résultat doit cependant être considéré comme un artefact de l'archive LTHE.

¹Le vent géostrophique est défini comme le vent que l'on observerait si les termes d'inertie et de frottements, notamment ceux liés au sol, étaient négligeables. C'est donc une sorte de vent idéal, qui se déduit directement des gradients de géopotentiels et qui se rapproche du vent réel dans les couches élevées de la troposphère. Ainsi, même si le vent géostrophique n'est pas le vent réel, on peut le considérer comme son *moteur*.

En effet, les variables définies sur des niveaux isentropiques ne sont disponibles qu'à des niveaux relativement élevés et à 0hTU (cf. 2.1.3.2). Or, comme nous le verrons plus loin, les variables mesurées à 12hTU revêtent une importance toute particulière pour la qualité de l'analogie vis-à-vis de notre prédictand.

Enfin, la réduction de la taille de la fenêtre d'analogie entraîne une diminution des performances sur toutes les variables. Elle tend en outre à favoriser le score $RMSE$ sur le critère $S1$:

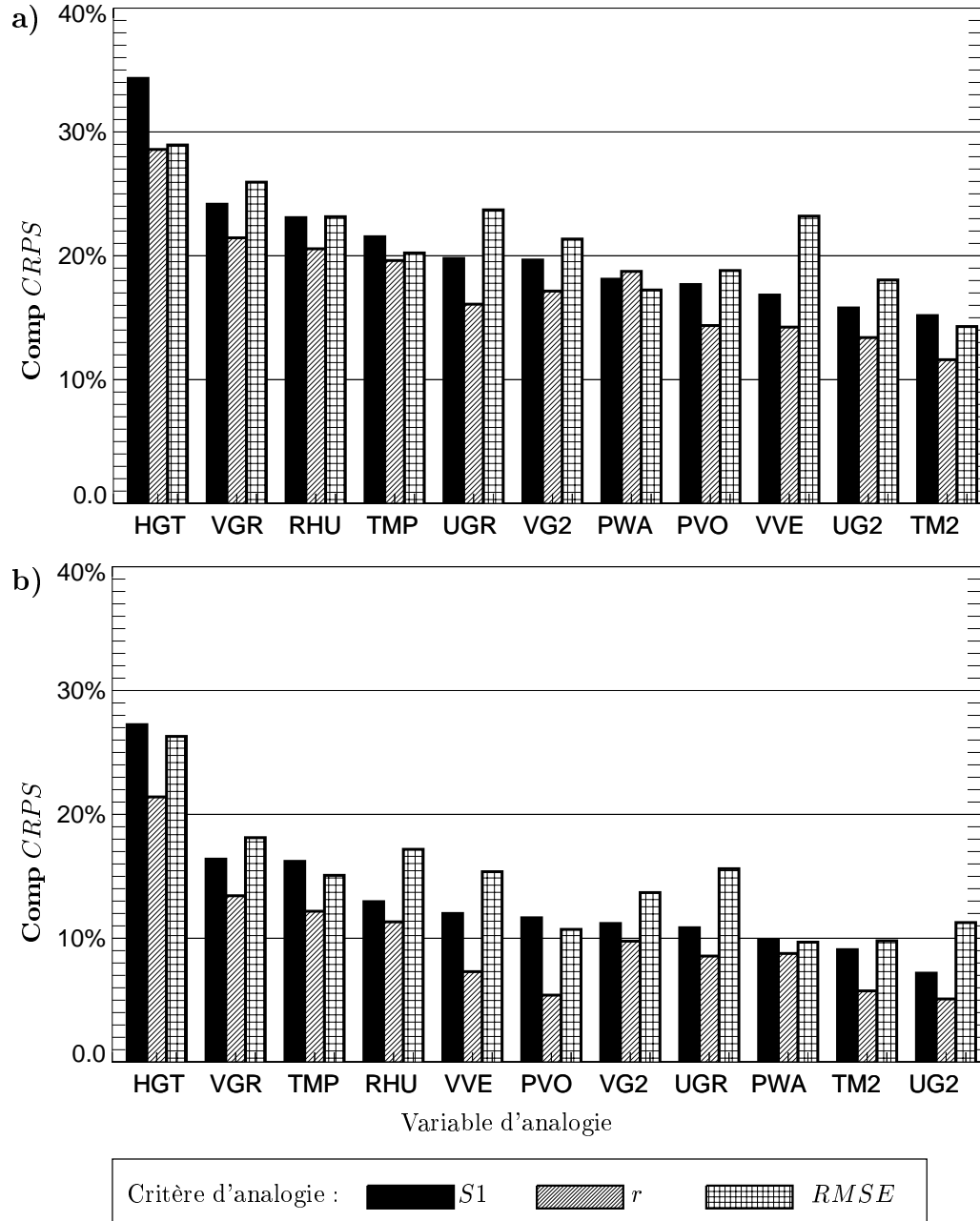


FIG. 4.5: Comparaison des performances moyennes sur l'ensemble des groupements, obtenues avec chaque type de variable disponible dans l'archive LTHE. Fenêtre d'analogie utilisée : **a)** fenêtre de la méthode de référence, **b)** fenêtre de $2.5^\circ \times 2.5^\circ$ centrée sur chaque groupement.

sur seulement 4 points de grille, la notion de structure spatiale d'un champ n'a plus vraiment de sens.

En conclusion, la prévision des précipitations journalières par analogie requiert en premier lieu une analogie de la circulation synoptique (sur une fenêtre adaptée), entre la situation cible et les situations candidates.

4.2 Utilisation d'un unique champ de géopotentiels

Un algorithme d'analogie performant semble devoir se baser sur la comparaison de champs de hauteurs géopotentielle, effectuée à l'aide du critère $S1$. Nous allons donc maintenant essayer d'aller plus loin dans cette direction en recherchant à quel niveau et à quel moment il est le plus judicieux de considérer ce champ géopotentiel.

Dans toute cette section nous nous fixons les conditions suivantes :

- **fenêtre d'analogie** : fenêtre de la méthode de référence,
- **critère d'analogie** : critère $S1$,
- **nombre d'analogues retenues** : 50 (méthode de référence).

4.2.1 Performances moyennes sur l'ensemble des groupements

Dans un premier temps, nous allons considérer une performance globale sur l'ensemble des groupements de l'archive EDF, en effectuant la moyenne arithmétique des performances atteintes sur chacun d'entre eux.

Cet ensemble de groupement est toutefois éminemment arbitraire. Leur seul point commun est de concerner, peu ou prou, des zones de montagne. Nous allons ainsi probablement lisser certains résultats, voire amoindrir la discrimination maximale souhaitable. Par contre, nous espérons faire apparaître les points communs et mettre ainsi en avant les variables d'analogies les plus générales possibles.

4.2.1.1 Toutes situations confondues

La figure 4.6a présente les valeurs de performances obtenues sur les prévisions de précipitations en fonction à la fois :

- du niveau (axe des ordonnées) du champ de hauteurs géopotentielle servant de support à l'analogie,
- et de l'heure de mesure (axe des abscisses) de ce même champ.

L'heure de mesure est exprimée de façon relative par rapport à l'instant 0 hTU de la journée J considérée, et est identique pour les journées cibles et les journées candidates. Ainsi, si pour comparer une journée cible C et une journée candidate J_{ana} on considère un champ à $+36h$, cela signifie que l'on va comparer le champ mesuré à 12hTU le jour $C + 1$ avec celui mesuré à 12hTU le jour $J_{ana} + 1$.

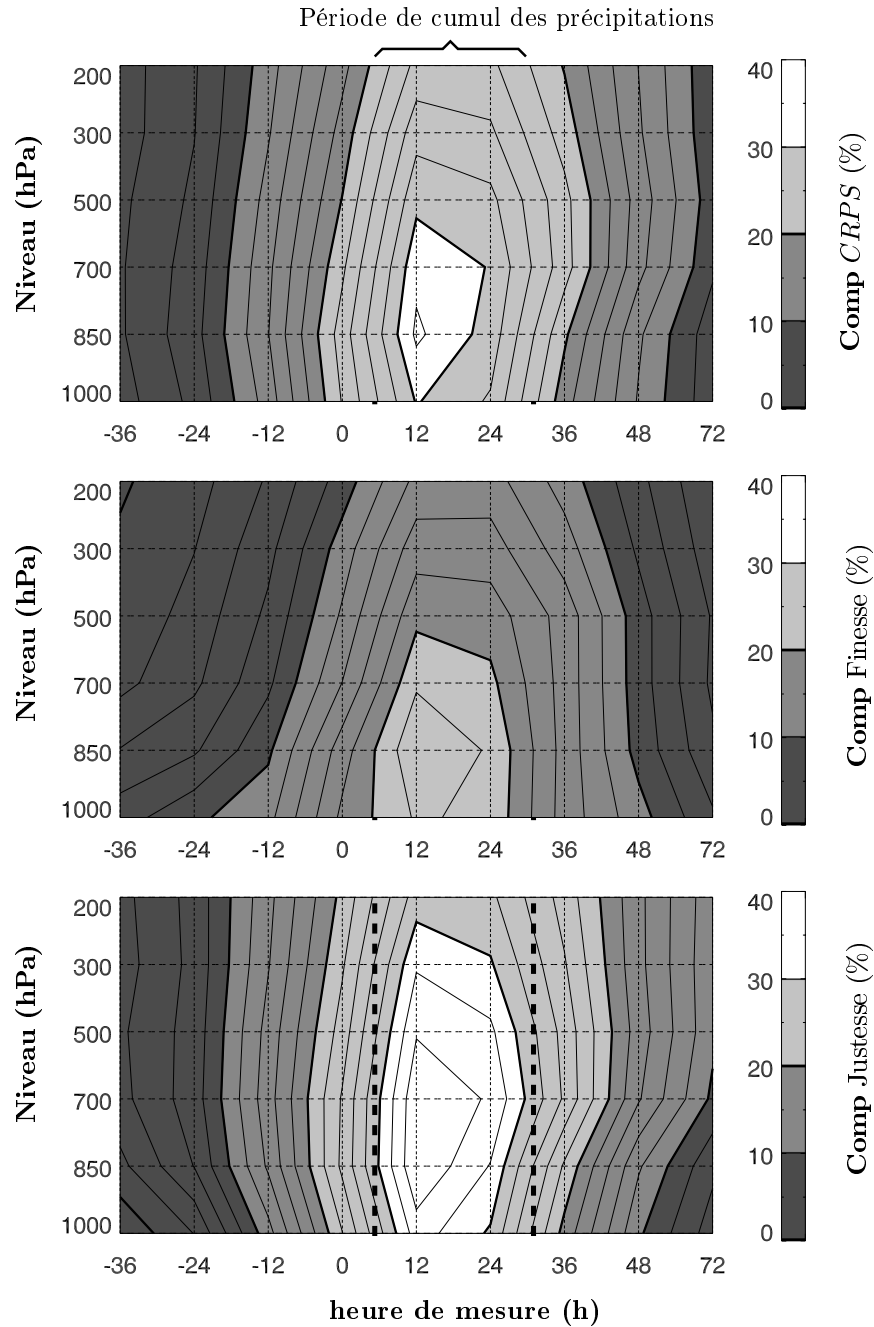


FIG. 4.6: Performances moyennes sur l'ensemble des groupements en fonction du niveau et de l'heure de mesure du champ de hauteurs géopotentiels. **a)** selon **Comp CRPS**; **b)** selon la compétence de la Finesse du *CRPS*; **c)** selon la compétence de la Justesse du *CRPS*.

Quelle que soit l'heure de mesure utilisée, les cumuls de pluie considérés restent ceux des journées C et J_{ana} . Nous rappelons que ces cumuls s'étalent en fait de 6hTU le jour C à 6hTU le jour $C + 1$ et de 6hTU le jour J_{ana} à 6hTU le jour $J_{ana} + 1$, ce qui correspond à des heures de mesure allant de +6h à +30h.

Les résultats montrent que l'heure de mesure du champ importe plus que son niveau : il semble nécessaire que la mesure soit effectuée à l'intérieur de la période de cumul des précipitations. Cela peut paraître évident a posteriori, mais montre néanmoins l'intérêt de disposer de mesures

météorologiques à un pas de temps au moins égal à cette période de cumul. Il est dommage que les mesures météorologiques à +18h ne soient pas disponibles au sein de l'archive LTHE : il est fort probable qu'elles auraient été les plus informatives¹.

Si l'heure de mesure s'écarte de cette plage optimale, les performances décroissent très vite. Ainsi, si l'on se place à -18h, soit 24 heures avant le début de la période de cumul, les performances sont réduites de plus de moitié. On confirme donc qu'un tel algorithme d'analogie, même s'il donne des résultats intéressants en *adaptation* (liaison entre une situation synoptique donnée et les précipitations concomitantes), est très vite dépassé en *prévision* (liaison entre une situation donnée et les précipitations futures).

Parmi les champs mesurés à +12 ou +24h, l'influence du choix du niveau considéré est plus perceptible. Les niveaux les plus intéressants se situent aux alentours de 850 à 700 hPa. Or, il se trouve que le champ de géopotentiel à 700 hPa est généralement considéré comme donnant une mesure raisonnable de la circulation de la moyenne troposphère : c'est celui-ci qui est utilisé traditionnellement dans la production des prévisions saisonnières au NCEP par exemple (Anderson *et al.* 1999).

En décomposant le *CRPS* en Finesse et Justesse, nous pouvons par ailleurs exprimer les performances obtenues en terme de Compétence de la Finesse et de Compétence de la Justesse (figures 4.6b et c). Il apparaît ainsi que la dépendance de la qualité des prévisions avec l'heure de mesure du champ utilisé est principalement liée à leur Justesse.

En effet, même si l'on considère un champ à -36h, les prévisions conservent plus de Finesse que la climatologie, alors qu'elles présentent une Justesse moindre. Ce comportement peut s'expliquer par le temps de persistance moyen d'un anticyclone, supérieur à 36 heures :

- Si les conditions sont anticycloniques à -36h, alors elles le sont aussi souvent à +6h. La prévision est donc caractérisée par une fréquence des valeurs nulles importante, ce qui lui donne une Finesse supérieure à celle de la climatologie. Par contre, elle ne montre pas forcément plus de Justesse en moyenne. En effet, si les conditions cibles à +6h ne s'avèrent pas anticycloniques et produisent au final des précipitations, la prévision sera fine mais en même temps très éloignée des observations, ce qui est fortement pénalisé par le *CRPS* (cf. 3.3.5.2).
- A l'inverse, si les conditions à +36h ne sont pas anticycloniques, elles ne préjugent en rien de ce qui va se dérouler dans plus de 42 heures. La prévision se résume donc alors à un échantillonnage aléatoire de la climatologie : sa compétence, aussi bien en terme de Finesse que de Justesse, est donc nulle.

4.2.1.2 Prévision des pluies intenses

En moyenne, la fréquence climatologique des valeurs nulles sur nos groupements est de l'ordre de 50%. Or, dans la plupart des applications hydrologiques de prévisions quantitatives de pré-

¹Les mesures à +18h sont produites par les réanalyses NCEP/NCAR mais ne sont malheureusement pas disponibles sur les CD-Rom annuels.

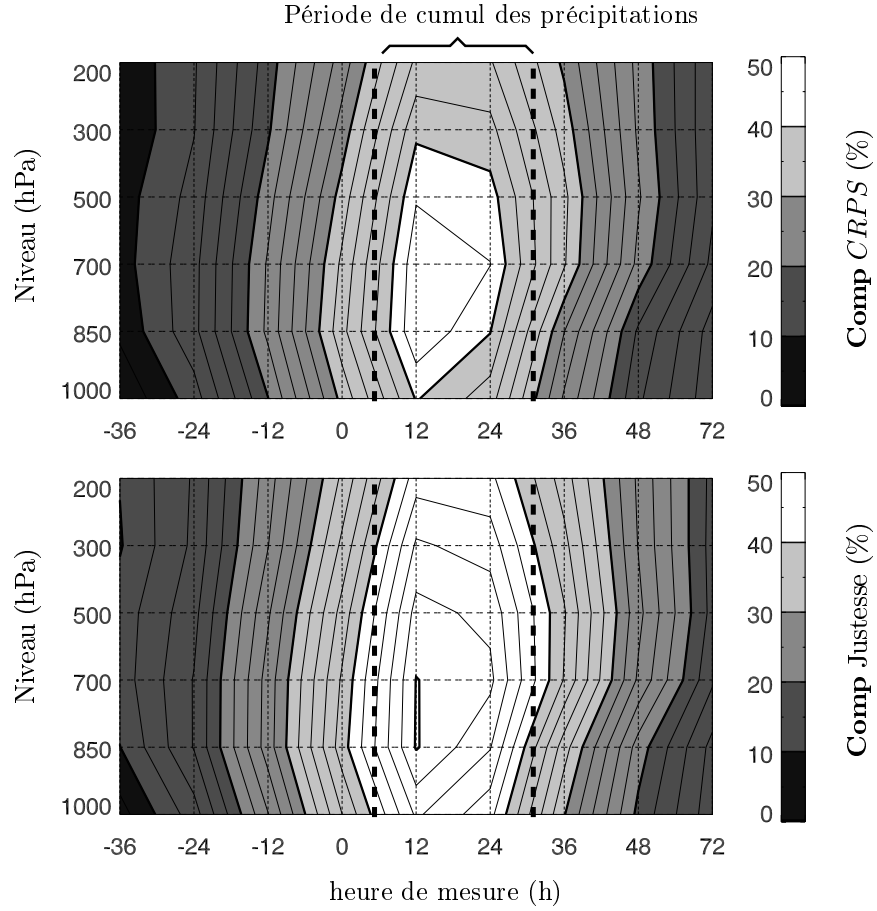


FIG. 4.7: Performances moyennes sur l'ensemble des groupements en fonction du niveau et de l'heure de mesure du champ de hauteurs géopotentiellelles, pour les situations cibles telles que $R \geq 0.5$. **a)** selon **Comp CRPS** ; **b)** selon la compétence de la Justesse du **CRPS**

cipitations, la prévision correcte des pluies nulles est secondaire devant celle des pluies intenses. On peut donc redouter que les résultats précédents soient trop influencés par les situations à précipitations nulles : les champs reconnus comme étant les plus informatifs en moyenne pour l'ensemble des situations constituent-ils aussi le meilleur choix pour prévoir les situations générant des précipitations intenses ?

Nous étudions donc à nouveau la dépendance de la qualité des prévisions en fonction du niveau et de l'heure de mesure du géopotentiel utilisé comme variable d'analogie, *mais* en restreignant les situations cibles aux seules situations présentant des précipitations supérieures à un seuil arbitrairement choisi (cf. équation 2.2) :

$$R \geq 0.5 \quad \Leftrightarrow \quad R_{brutes} \geq \frac{R10}{4} \quad (4.5)$$

Les situations cibles sont donc différentes d'un groupement à l'autre. Les situations candidates restent par contre les mêmes, si bien que la climatologie de référence intervenant dans le calcul des scores de compétence est elle aussi identique.

Cependant, les résultats, présentés par la figure 4.7, montrent une étonnante concordance avec les résultats précédents. Que l'on s'intéresse aux fortes précipitations ou à toute la gamme des précipitations, le géopotentiel le plus informatif est le même, à savoir 850 hPa à +12h. Ceci est fort intéressant car nous disposons donc d'un algorithme général : il n'est pas besoin de donner une priorité à la prévision des forts cumuls ou à celle de l'absence de précipitations.

La distribution climatologique utilisée comme référence, qui possède une fréquence des valeurs nulles importante, fournit bien sûr une prévision très peu utile pour les fortes pluies. C'est pourquoi on peut noter que les valeurs de compétence obtenues sur les pluies intenses sont globalement supérieures à celles obtenues sur l'ensemble des précipitations. Toutefois, la distribution climatologique demeure particulièrement fine, alors que les distributions issues de l'adaptation, prévoyant beaucoup moins de valeurs nulles, le sont évidemment moins. La compétence de la Finisse est donc ici inférieure à 0 : présentant peu d'intérêt dans ce cas, elle n'a pas été représentée.

4.2.2 Influence du choix du groupement cible

4.2.2.1 Niveaux et heures de mesure optimaux

Après avoir considéré les groupements pluviométriques dans leur ensemble, il est intéressant de se pencher sur chacun d'entre eux pris séparément. Il sera ainsi possible de voir si le champ optimal est le même pour tous les groupements, et sinon, s'il suit une répartition spatiale cohérente. La fenêtre d'analogie reste toutefois inchangée (fenêtre de référence) et demeure la même pour tous.

Nous avons ainsi calculé pour chaque groupement, en considérant à nouveau l'ensemble des situations cibles, quels étaient les niveaux et heures de mesure optimaux du géopotentiel choisi comme variable d'analogie (figure 4.8).

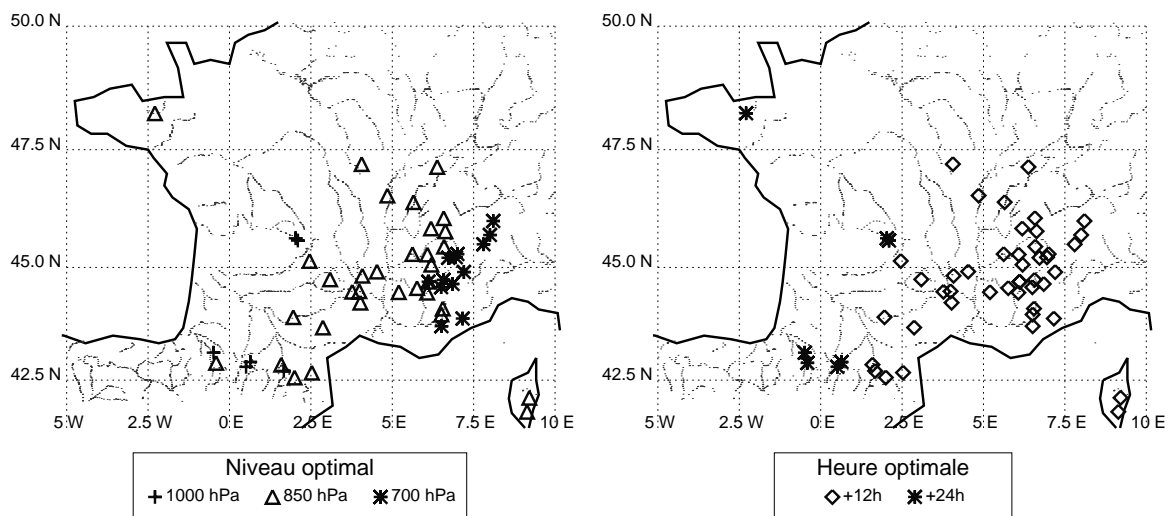


FIG. 4.8: Cartographie des niveaux et des heures de mesure optimaux en fonction du groupement pluviométrique considéré.

La grande majorité de nos groupements considère une heure de mesure optimale de +12h. Seuls les 7 groupements situés les plus à l'ouest préfèrent une mesure à +24h. On peut tenter d'interpréter ce décalage en considérant simplement une description lagrangienne du déplacement moyen des masses d'air sous nos latitudes. En effet, la fenêtre d'analogie utilisée est excentrée par rapport à ces 7 groupements. Les masses d'air se déplaçant globalement d'Ouest en Est, celles qui passent dans la fenêtre d'analogie à +24h, se trouvaient généralement plus ou moins à l'aplomb de ces 7 groupements à +12h (figure 4.9). Celles situées dans la fenêtre à +12h correspondent en fait à celles qui intéressaient ces groupements à +0h, ce qui est loin d'apporter les meilleures performances. Il semble donc que, quel que soit le groupement considéré, dès lors que la fenêtre d'analogie n'est pas trop excentrée par rapport à celui-ci, l'heure de mesure optimale du géopotentiel est la même pour tous, à savoir +12h.

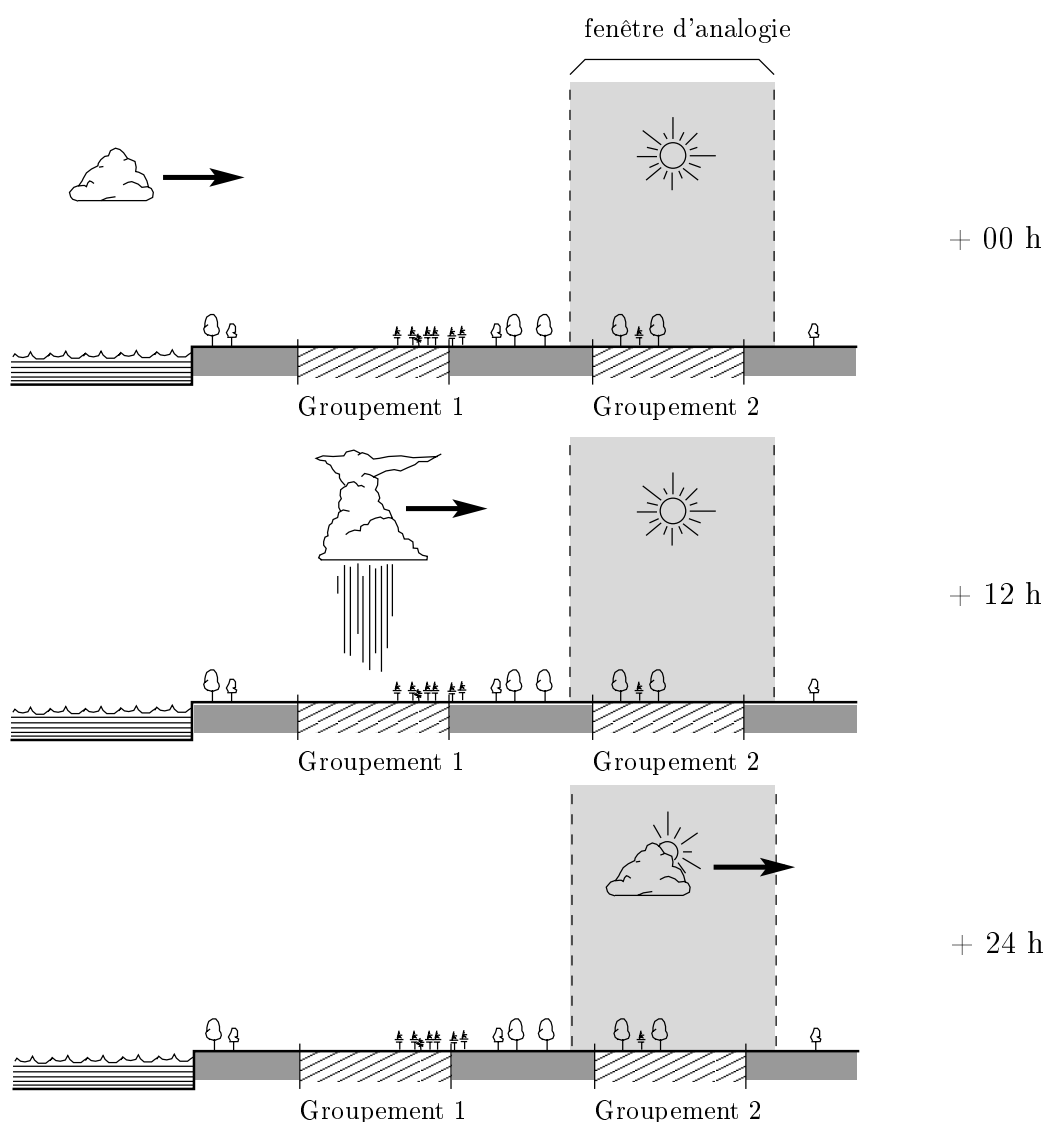


FIG. 4.9: Représentation schématique de la description lagrangienne du déplacement des masses d'air : la masse d'air présente dans la fenêtre d'analogie à +24h est celle qui concernait le groupement 1 à +12h.

La carte de répartition des niveaux optimaux est pour sa part plus délicate à interpréter. Elle semble être en partie liée à l'altitude moyenne des groupements. Ainsi, une grande partie des groupements alpins ont intérêt à utiliser un niveau plus élevé que la moyenne. Toutefois, à titre de contre-exemple, on peut citer le bassin *Hautes_Pyrénées_Centrales*, qui, bien que situé en altitude, préfère utiliser un niveau 1000 hPa.

4.2.2.2 Compétence des prévisions

La performance des prévisions émises dépend donc du choix du géopotentiel utilisé, mais elle est aussi très variable d'un groupement à l'autre. Les figures 4.10 et 4.11 présentent, groupement par groupement, les performances obtenues en considérant le géopotentiel 850 hPa à +12h.

La répartition spatiale des niveaux de performance semble s'organiser autour du centre de gravité des groupements. Plus on s'en éloigne, moins l'adaptation est efficace. Or, la fenêtre d'analogie utilisée a été optimisée par S. Guilbaud en considérant la performance moyenne de la prévision sur l'ensemble des bassins. C'est donc pour les bassins proches du centre de gravité qu'elle est la mieux adaptée. Nous verrons dans le chapitre suivant comment s'affranchir de cet effet en considérant plusieurs fenêtres d'analogie.

L'influence de la fenêtre n'explique néanmoins pas toute la variabilité que l'on observe. Ainsi, le groupement *Drôme* et ses voisins (*Verdon*, *Durance_moyenne* ...), bien que proches du centre de gravité, présentent des compétences relativement faibles. Par ailleurs, les groupements proches les uns des autres et soumis à des influences météorologiques comparables présentent des niveaux de performances similaires.

Une partie de ces disparités de qualité de prévision pourrait peut-être s'expliquer par la façon dont les groupements pluviométriques ont été créés. Faute par exemple de stations de mesure appropriées, certains peuvent être hétérogènes ou trop localisés. Ils sont en outre de tailles très diverses, ce qui influe sur leur potentiel de prévisibilité (cf. 2.2).

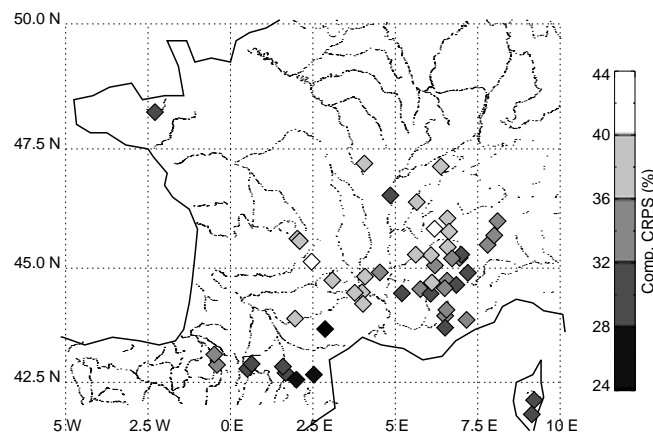


FIG. 4.10: Carte de la Compétence du CRPS (prévisions basées sur le géopotentiel 850 hPa à +12h).

Une autre partie semble pour sa part plutôt liée à des spécificités régionales : tous les groupements ne réagissent pas de la même façon aux conditions synoptiques. Si l'on considère la décomposition du *CRPS* en terme de Finesse et de Justesse (figure 4.11), on s'aperçoit que la compétence de la Finesse est fortement liée au paramètre $F0$ de chaque groupement (coefficient de détermination $R^2 = 0.81$). Ainsi, plus la fréquence climatologique des valeurs nulles $F0$ est forte, moins la compétence de la Finesse est importante. En effet, si $F0$ est grand, alors la dis-

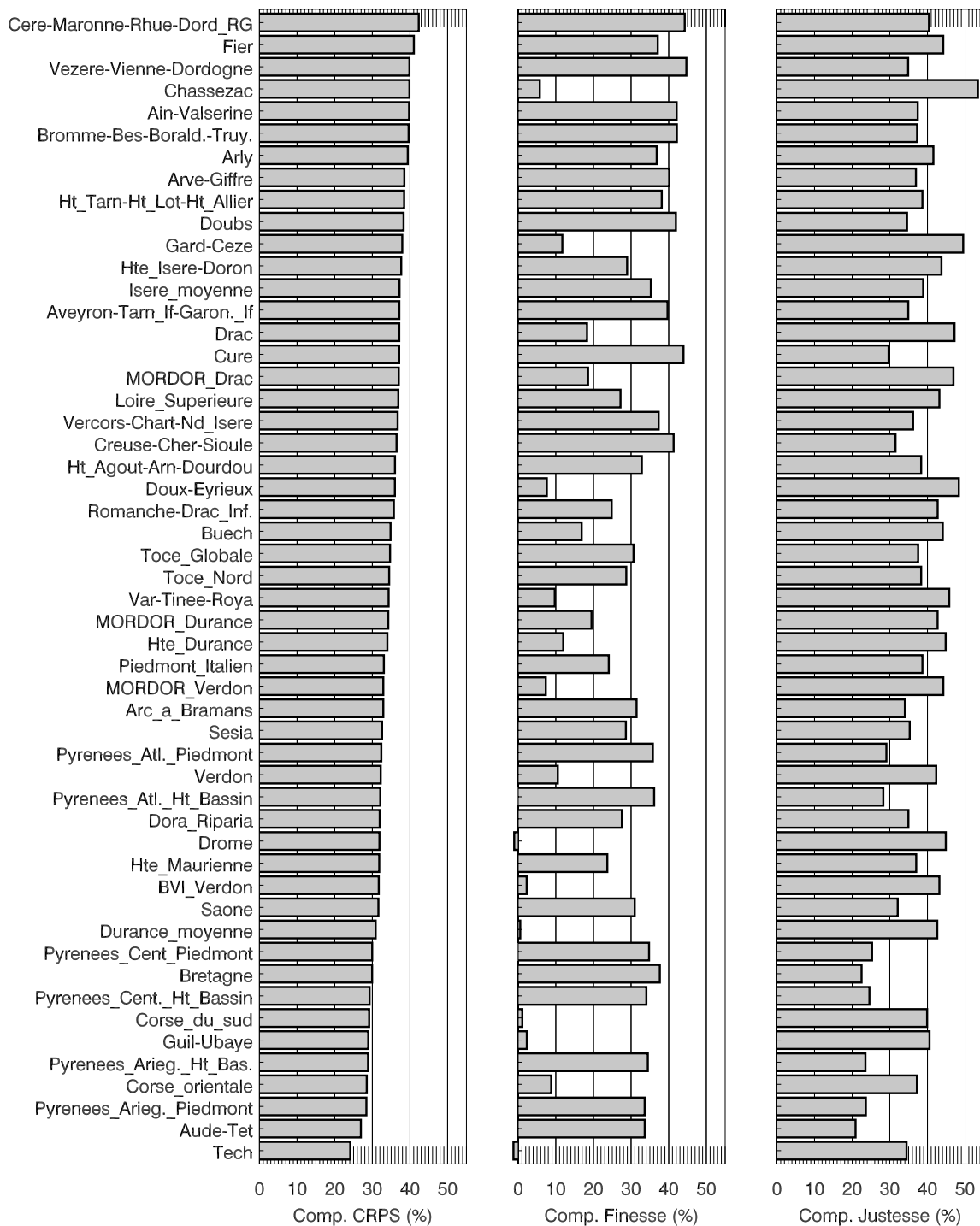


FIG. 4.11: Compétence du CRPS ainsi que de sa décomposition en Finesse et Justesse pour chaque groupement (prévisions basées sur le géopotentiel 850 hPa à +12h.)

tribution climatologique est forcément fine : pour une même Finesse de prévision, la compétence est inférieure.

Le *CRPS* et la Justesse ne présentent néanmoins pas d'aussi fortes liaisons avec *F0* : leurs variations ne semblent pas dues à un simple artefact de construction de ces scores. Pour tenter de les expliquer, il peut être intéressant de reprendre ici la décomposition développée précédemment (cf. 1.3.4), d'une précipitation $P_{obs}^{S^0}$ observée lors d'une situation météorologique S^0 , en :

- une partie $P_{syn}^{S^0}$ expliquée par les conditions synoptiques (ou plus exactement les variables d'analogie),
- et une autre partie $P_{locales}^{S^0}$ expliquée par les conditions locales (ou plus exactement les variables non prises en compte par l'analogie).

La partie expliquée par les conditions synoptiques dépend du régime météorologique auquel est soumis le groupement, ainsi qu'à sa morphologie. Elle est donc propre à chaque bassin mais peut être organisée de façon régionale.

Il est donc possible que les différences de performance observées entre les groupements, hormis pour les plus excentrés par rapport à la fenêtre d'analogie, soient d'avantage dues à leurs spécificités climatiques propres, qu'au choix de l'utilisation d'un même algorithme d'adaptation pour tous (même niveau, même heure de mesure du géopotential). Pour confirmer cette hypothèse, nous avons tracé, pour les deux groupements réalisant la meilleure et la moins bonne prévision (respectivement **Cère** et **Tech**), l'influence sur la compétence des prévisions du choix du niveau et de l'heure de mesure du géopotential servant de support à l'analogie (figure 4.12).

Il apparaît que, à un facteur d'échelle près, l'organisation des isolignes obtenues est effectivement très semblable. Une analogie sur la circulation générale est donc reliée de la même façon aux précipitations sur les deux groupements, mais l'importance relative de cette liaison n'est pas la même. Le groupement **Cère**, situé sur le flanc Est du Massif Central, est principalement intéressé

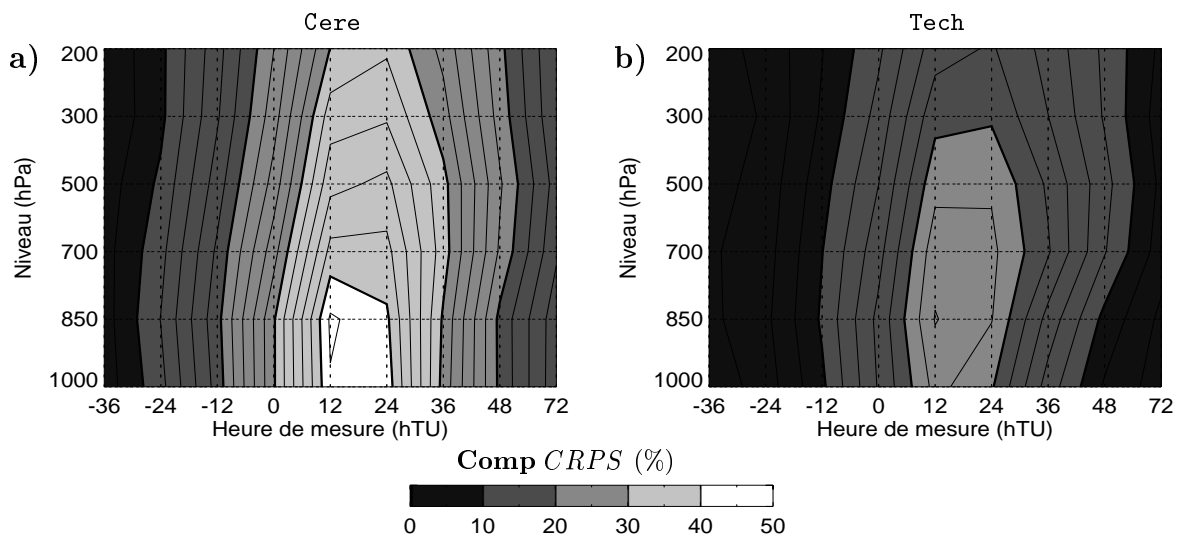


FIG. 4.12: Compétence du *CRPS* en fonction du niveau et de l'heure de mesure du géopotential, pour **a)** le groupement **Cère**, et **b)** le groupement **Tech**.

par les épisodes frontaux venant de l'Atlantique, fortement reliés à la circulation synoptique. Le groupement **Tech**, situé pour sa part au pied des Pyrénées, sur leur versant Est, se trouve plus concerné par des événements convectifs venant buter contre les Pyrénées depuis la Méditerranée. Ces événements se déroulent à des échelles spatiales inférieures à celle de la circulation synoptique et sont donc moins bien expliqués par notre analogie.

4.3 Combinaison de plusieurs champs de géopotentiels

Nous n'avons considéré jusqu'à présent que des analogies basées sur un seul champ de hauteurs géopotentielle. Or, il peut être judicieux de prendre en compte plusieurs géopotentiels, mesurés à des niveaux et des heures différents afin d'enrichir l'analogie. En effet, en disposant d'une description plus complète de la circulation atmosphérique, on peut espérer obtenir une liaison plus significative entre l'analogie et les précipitations.

4.3.1 Comment prendre en compte plusieurs champs ?

Pour considérer une analogie basée sur plusieurs géopotentiels, nous avons opté pour la même solution que celle adoptée par S. Guilbaud (1997) : la distance d'analogie DA , qui permet de classer entre elles les situations candidates $\{J_{ana}\}$ par rapport à la journée cible C , est obtenue en faisant simplement la moyenne arithmétique des scores $S1$ calculés sur chacun des champs de l'analogie :

$$DA(J_{ana}) = \frac{1}{N_{Ch}} \sum_{i_{Ch}=1}^{N_{Ch}} S1\left(HGT_{i_{Ch}}^C; HGT_{i_{Ch}}^{J_{ana}}\right) \quad (4.6)$$

où : $\left\{ \begin{array}{l} N_{Ch} \text{ est le nombre de géopotentiels intervenant comme variable d'analogie,} \\ HGT_i^J \text{ est le } i^{\text{ème}} \text{ géopotential des } N_{Ch} \text{ géopotentiels intervenant comme variable} \\ \text{d'analogie, de la situation } J. \end{array} \right.$

Cependant, des scores calculés pour des niveaux différents sont-ils réellement comparables entre eux ? Un géopotential situé près du sol a effectivement une texture différente de celle d'un géopotential situé en altitude. Ce dernier est généralement plus lisse et présente un gradient Nord-Sud plus systématique. Il est donc plus facile d'obtenir des valeurs de $S1$ faibles quand on s'intéresse à des champs élevés, comme le montre la figure 4.13. Ainsi, selon le niveau considéré, un même gain de $S1$ n'a pas forcément la même signification.

Nous nous sommes donc demandé s'il n'était pas plus rigoureux de normaliser les $\{S1_{i_{Ch}}\}$ avant de les moyenner en une distance d'analogie. Nous définissons ainsi le score normalisé $S1^*$ comme la valeur centrée réduite de $S1$ par rapport aux distributions présentées figure 4.13 et résumées dans le tableau 4.1 :

$$S1_{i_{Ch}}^* = \frac{S1_{i_{Ch}} - m_{i_{Ch}}}{\sigma_{i_{Ch}}} \quad (4.7)$$

où :

- $m_{i_{Ch}}$ est la moyenne sur l'archive du $i_{Ch}^{\text{ème}}$ géopotentiel intervenant comme variable d'analogie,
- $\sigma_{i_{Ch}}$ est l'écart-type sur l'archive du $i_{Ch}^{\text{ème}}$ géopotentiel intervenant comme variable d'analogie.

Nous pouvons ainsi calculer une distance d'analogie normalisée DA^* à partir de la moyenne des scores normalisés $S1_{i_{Ch}}^*$.

La dimension du domaine spatial sur lequel est calculé le score $S1$ n'est bien sûr pas sans influence sur les caractéristiques de ces distributions. Lorsque le domaine augmente, on voit que les distributions ont tendance à se raidir : le nombre de degrés de liberté augmentant, il est plus

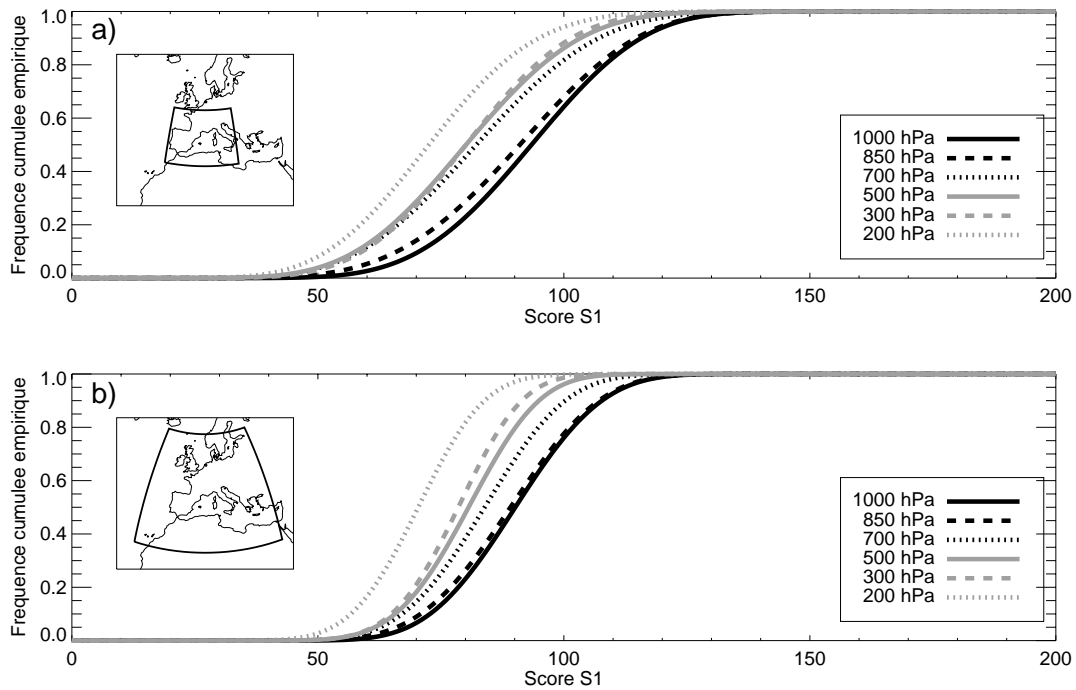


FIG. 4.13: Fonctions de répartition des scores $S1$ en fonction du niveau du géopotentiel considéré. Ces résultats concernent, sur l'ensemble de la période 1953-96, tous les couples de situations qui se trouvent à plus ou moins 2 mois près dans la même période de l'année. La zone géographique sur laquelle est calculée le score $S1$ est : **a)** $-10/17.5^\circ\text{E } 32.5/50^\circ\text{N}$, et **b)** $-20/32.5^\circ\text{E } 25/62.5^\circ\text{N}$.

		Niveau de géopotentiel					
		1000 hPa	850 hPa	700 hPa	500 hPa	300 hPa	200 hPa
Grille	m	93	90	82	80	80	73
-10/17.5°E 32.5/50°N	σ	18.7	21.7	20.9	19.1	17.4	16.8
Grille	m	90	89	83	80	78	70
-20/32.5°E 25/62.5°N	σ	14.7	14.3	17.2	14.1	12.6	12.6

TAB. 4.1: Moyennes (m) et écart-types (σ) des distributions cumulées empiriques des scores $S1$ en fonction du niveau du champ de hauteurs géopotentielle considéré.

difficile de trouver des situations soit très ressemblantes, soit à l'inverse en totale opposition de phase.

Il s'avère cependant que, d'un niveau à l'autre, les écart-types $\sigma_{i_{Ch}}$ sont peu variables. La principale différence entre les distributions est due à leurs écarts de moyenne $m_{i_{Ch}}$. Or, ces écarts sont sans incidence sur le classement relatif des DA ou des DA^* ; seules comptent les différences d'écart-type. Ainsi, si l'on considère que tous les écart-types sont égaux à σ , nous avons :

$$DA^* = \frac{DA}{\sigma} - \frac{1}{N_{Ch}} \cdot \frac{\sum_{i_{Ch}}^{N_{Ch}} m_{i_{Ch}}}{\sigma} \quad (4.8)$$

Donc, pour deux situations J_1 et J_2 candidates à représenter une même situation cible :

$$DA^*(J_1) \leq DA^*(J_2) \quad \Leftrightarrow \quad DA(J_1) \leq DA(J_2) \quad (4.9)$$

Ainsi, si les écart-types de score $S1$ des différents niveaux sont tous égaux, le classement des journées candidates est identique, que la distance d'analogie soit normalisée ou non.

Dans le cas présent, les écart-types sont effectivement très proches les uns des autres. Différents essais d'applications de la distance d'analogie normalisée ont confirmé que la normalisation était une précaution inutile ici. Nous nous en tiendrons donc à la simple moyenne.

Pour mémoire, S. Guilbaud avait essayé d'introduire des pondérations différenciées sur les différents niveaux, en les faisant varier de manière heuristique autour de 1. Elle cherchait ainsi à voir s'il fallait donner plus de poids aux niveau sol ou altitude. Cette tentative n'avait pas non plus été concluante.

4.3.2 Les meilleures combinaisons !

4.3.2.1 Utilisation de deux champs de hauteurs géopotentielle

Pour trouver le couple de géopotentiels le plus informatif en matière de prévision de précipitations par analogie, nous avons passé en revue toutes les combinaisons possibles, en faisant varier les heures de mesure de -12 à +48 et en considérant les 6 niveaux disponibles (ce qui fait en tout 1260 couples possibles). La fenêtre d'analogie utilisée est une fenêtre optimisée pour l'utilisation du géopotentiel le plus pertinent : 850 hPa à +12h (cf. 5.1 pour plus de détails sur l'optimisation de la fenêtre).

Les performances obtenues sur l'ensemble des groupements par les 50 meilleurs couples sont présentées par la figure 4.14. On constate immédiatement que pour ceux-ci, l'utilité des prévisions de précipitation est effectivement accrue lorsque l'on considère non plus un seul, mais deux champs de hauteurs géopotentielle. Il existe néanmoins un grand nombre de couples qui présentent des performances moindres ; la richesse de l'information n'est pas suffisante si elle n'est pas pertinente. Ainsi, adjoindre le géopotentiel 1000 hPa à +0h (noté z1000+0) au géopotentiel 850 hPa à +12h (z850+12) dégrade les performances de l'analogie. En effet, les situations

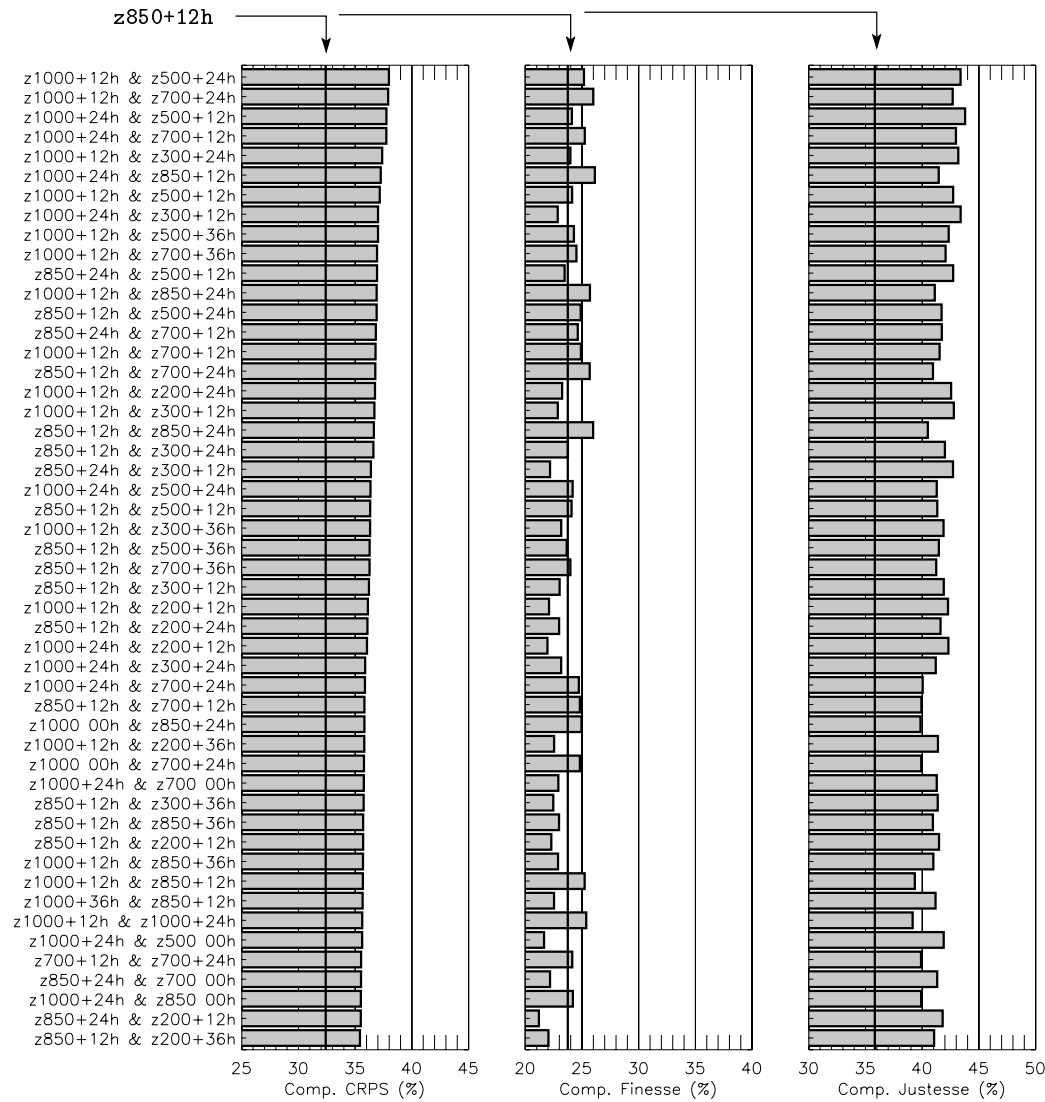


FIG. 4.14: Classement des 50 couples de géopotentiels les plus informatifs (dans le libellé des couples, z1000+12h correspond au géopotentiel situé à 1000 hPa et mesuré à +12h)

analogues doivent ressembler à la fois à z1000+0 et à z850+12, ce qui exclut les situations qui, même si elles sont fortement analogues à z850+12, ne présentent pas de similitudes à z1000+0. Or, la ressemblance à z1000+0 est finalement assez peu pertinente comme nous l'avons vu précédemment. Au final, on « dilue » ainsi l'information utile contenue dans z850+12.

Le choix du couple de géopotentiel servant de support à l'analogie ne se fait donc pas au hasard. En scrutant plus en détail les résultats obtenus, on voit apparaître certaines règles :

- Les couples qui donnent les meilleurs résultats font intervenir :
 - > à la fois un géopotentiel mesuré à +12h,
 - > et un autre mesuré à +24h,
 confirmant ainsi l'importance des heures de mesure comprises dans la période de cumul.
- Ils prennent de même en compte :
 - > à la fois un géopotentiel mesuré dans les basses couches : 1000 ou 850 hPa, qui reflète plutôt les conditions locales des écoulements atmosphériques,
 - > et un autre mesuré en altitude : 750, 500 ou 300 hPa, qui représente les grandes lignes de la circulation atmosphérique générale.

Par contre, le fait de mesurer le géopotentiel en basse couche à +12h ou à +24h ne semble pas avoir une grande importance, hormis peut-être un léger gain en finesse à +12h. Ce qui apparaît comme primordial, c'est bien de combiner des géopotentiels à des niveaux et des heures de mesure différents en respectant les deux règles empiriques exprimées ci-dessus, à l'image du couple le plus performant : $z1000+12$ et $z500+24$.

Comme le montre la figure 4.14, le gain de performance est essentiellement dû à un gain de Justesse de la prévision, la Finesse ayant même tendance à diminuer. Les journées cibles pluvieuses reconnaissent en effet moins de journées sèches comme analogues, si bien que les distributions prévues possèdent un point d'accumulation en zéro moins prononcé. Elles présentent donc une Finesse moindre alors leur Justesse augmente.

Il est par ailleurs possible de considérer l'évolution des performances obtenues en faisant varier l'un des champs de ce couple, tout en gardant l'autre fixe (figure 4.15). On constate alors quelques comportements intéressants :

- Il y a une différenciation assez nette entre l'intérêt des champs de basses couches et celui des champs d'altitude, qui apparaissent comme complémentaires. La transition d'un domaine à l'autre semble s'effectuer entre 850 et 700 hPa, c'est-à-dire plus ou moins vers le sommet de la couche limite atmosphérique continentale.
- Une fois que le champ de basse couche est fixé, l'optimum pour le second champ est assez peu marqué : considérer $z300+24$ ou $z750+24$ à la place de $z500+24$ ne fait pas chuter les performances de façon dramatique. A l'inverse, si l'on fixe le champ d'altitude, on restreint fortement les possibilités pour le champ de basse couche.

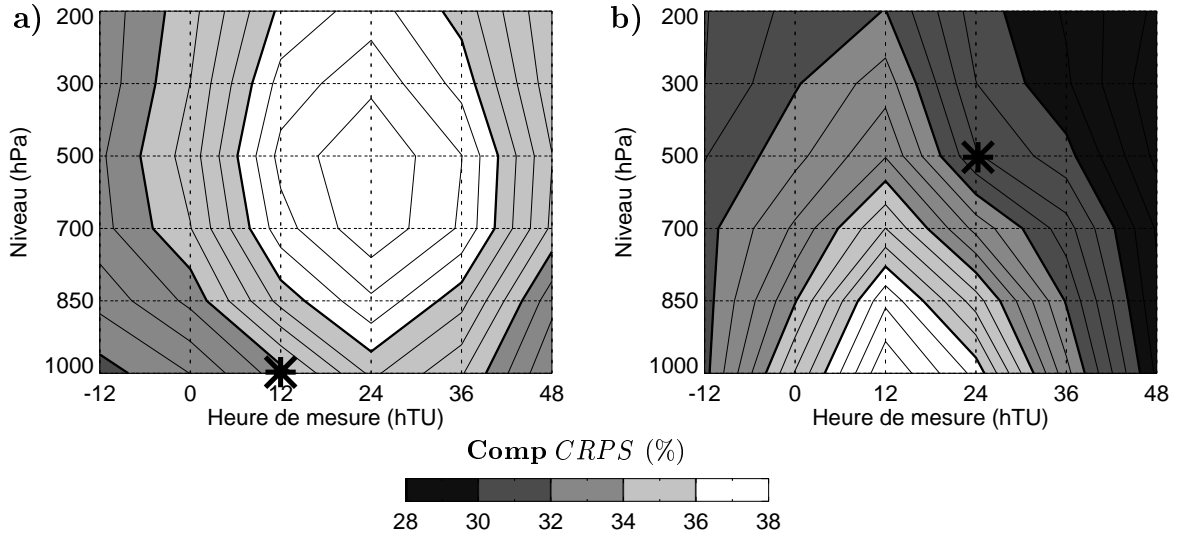


FIG. 4.15: Compétence du *CRPS* sur l'ensemble des groupements en basant l'analogie sur un couple de géopotentiels dont l'un est fixé, en fonction du niveau et de l'heure de mesure de l'autre. L'étoile sur chaque graphique représente le géopotentiel fixé : **a)** 1000 hPa à +12h, et **b)** 500 hPa à +24h.

4.3.2.2 Utilisation de 3 et 4 champs

L'utilisation combinée de deux champs de hauteurs géopotentielles différents comme variable d'analogie s'étant montrée fructueuse, nous avons étendu nos essais en combinant 3 puis 4 champs.

Toutefois, le nombre de combinaisons possibles avec trois champs devient prohibitif et nous n'avons pas pu les tester toutes de façon exhaustive. Afin de garder des temps de calcul raisonnables, nous n'avons essayé que les combinaisons dans lesquelles se trouve l'un des 25 meilleurs couples identifiés précédemment. De même, en ce qui concerne les combinaisons à 4 champs, nous n'avons considéré que celles qui contiennent l'un des 25 meilleurs triplets.

Bien que cet aspect ne soit décrit qu'au chapitre suivant, nous précisons que la fenêtre d'analogie utilisée pour comparer ces combinaisons entre elles est :

- celle optimisée sur le meilleur couple pour les combinaisons à 3 champs,
- et celle optimisée sur le meilleur triplet pour celles à 4 champs.

Cependant, les résultats obtenus sont présentés sur la figure 4.16 en utilisant, pour chacune des combinaisons, la fenêtre d'analogie qui lui est optimale. Ainsi, chaque combinaison présente son maximum possible de performance.

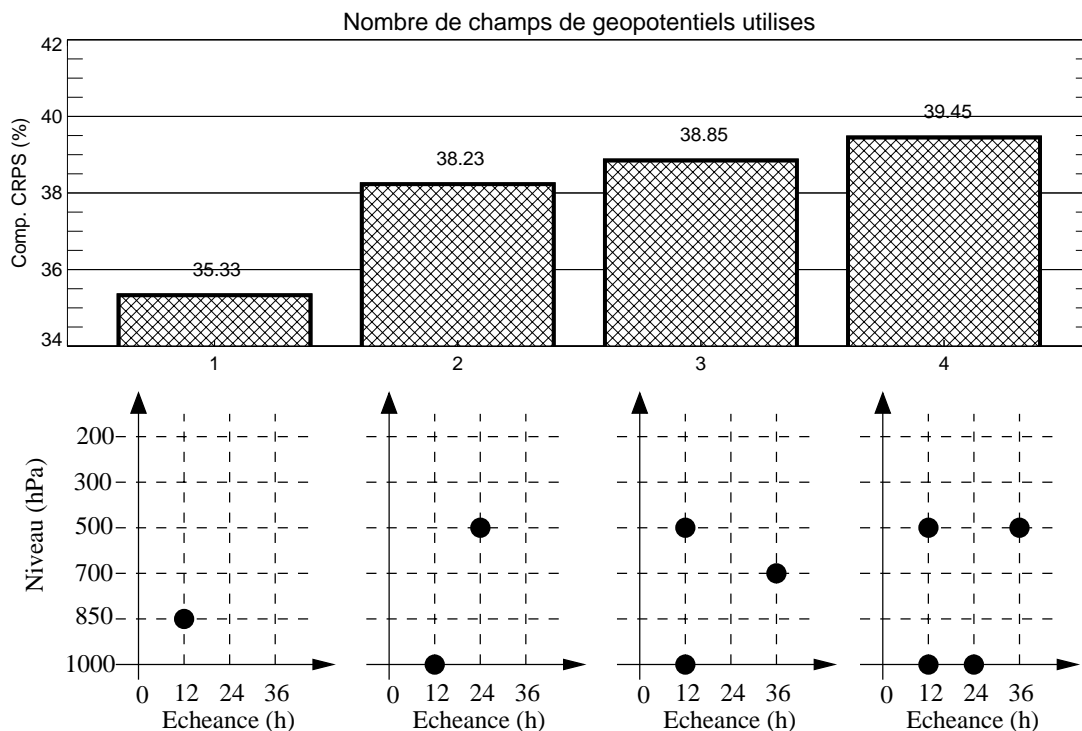


FIG. 4.16: Evolution de la compétence du *CRPS* en fonction du nombre de géopotentiels utilisé, en considérant à chaque fois la meilleure combinaison possible (présentée dans le panneau inférieur).

Les meilleures combinaisons s'avèrent être (z1000+12, z500+12, z700+36) pour 3 champs, et (z1000+12, z500+12, z1000+24, z500+36) pour 4 champs. Ces résultats montrent encore la complémentarité entre champ de basse couche et champ d'altitude.

Les gains obtenus en considérant plus que deux champs demeurent néanmoins faibles par rapport à ceux réalisés en passant de un à deux champs. On peut par ailleurs être étonné en

voyant apparaître un champs mesuré à +36h dans les meilleures combinaisons, alors que cette heure de mesure se trouve hors de la période de cumul des précipitations. Ceci peut peut-être s'expliquer par le caractère excentré de la fenêtre d'analogie par rapport aux groupements situés le plus à l'Ouest (cf. 4.2.2.1).

En outre, ces champs à +36h apportent un nouveau problème lorsque l'on se place dans l'optique d'une application opérationnelle en adaptation. Dans ce cas, les géopotentiels sont en effet prévus par un modèle NWP. Or, plus l'échéance de prévision est éloignée, moins ces prévisions sont fiables : les champs à +36h sont de moins bonne qualité que ceux à +24h, qui sont eux-mêmes moins fiables que ceux à +12h. Le gain apporté en prévision parfaite par la prise en compte d'un deuxième géopotentiel semble justifier le fait d'utiliser un champ à +24h en plus de celui à +12h, mais le léger gain obtenu avec un troisième champ ne risque-t-il pas d'être compensé (voire de se transformer en perte) par la moins bonne qualité du champ à +36h ?

L'adjonction d'un troisième ou quatrième champ semble donc ne pas être réellement intéressante : elle ne fait que légèrement corriger le choix de la fenêtre d'analogie, tout en alourdissant les algorithmes et en rendant opérationnellement l'adaptation plus sensible aux performances du modèle NWP utilisé pour prévoir les champs de géopotentiels.

4.4 Conclusions

Nous avons dans un premier temps défini plusieurs critères d'analogie, afin de tester la pertinence des différentes variables présentes dans l'archive LTRE pour une utilisation en adaptation par analogie. Ce test préliminaire a montré que, quelle que soit la taille de la fenêtre d'analogie considérée, les meilleurs résultats sont obtenus en comparant des champs de hauteurs géopotentiels via le critère d'analogie *S1*. Celui-ci s'attache à comparer les gradients de ces champs, si bien que l'analogie porte au final sur les vents géostrophiques et donc sur la circulation atmosphérique, qui s'est ainsi révélée être la meilleure variable d'analogie vis-à-vis des précipitations.

Pour obtenir les meilleurs résultats, il est par ailleurs important de bien choisir le géopotentiel sur lequel est basée l'analogie. Il s'est avéré que celui-ci devait être mesuré pendant la période de cumul des précipitations, c'est à dire dans le cas présent, à +12 ou +24h. Il doit en outre se situer à un niveau intermédiaire entre les basses couches de l'atmosphère et l'atmosphère totalement libre : entre 850 et 700 hPa. Finalement, en considérant les données dont nous disposons et nos objectifs de prévision, le géopotentiel le plus pertinent est $z850+12$.

Si l'on fait porter l'analogie sur une combinaison de deux géopotentiels et non plus sur un seul, on voit les performances augmenter de façon significative. L'un des champs choisis doit alors être mesuré dans les basses couches tandis que l'autre doit l'être en altitude. Leurs heures de mesure doivent, de préférence, être décalées et se situer à nouveau dans la période de cumul des précipitations. La meilleure combinaison est pour nous ($z1000+12$, $z500+24$).

La prise en compte d'un nombre supérieur de géopotentiels apporte des améliorations qui ne paraissent pas réellement significatives devant la complexification des algorithmes et les difficultés d'interprétation que cela entraîne pour le prévisionniste.

Enfin, il est apparu lors de ce chapitre, que même si les géopotentiels pertinents sont globalement identiques pour tous les groupements, les performances obtenues sont très variables d'un groupement à l'autre. La partie des précipitations que l'on peut expliquer par l'intermédiaire de la circulation synoptique semble en effet dépendre des influences climatiques propres à chaque groupement, voire de la façon dont le groupement a été constitué (taille, nombre et représentativité des stations de mesure, ...).

5

Du choix de la fenêtre d'analogie

Lors du chapitre précédent, le choix de la meilleure variable d'analogie a été longuement discuté. Maintenant, il reste en suspens la question du domaine spatial sur lequel celle-ci doit être considérée : la fenêtre d'analogie.

Nous allons tâcher de répondre à cette question en considérant successivement deux contextes différents. Nous utiliserons ainsi tout d'abord des fenêtres définies une fois pour toutes et qui restent ensuite les mêmes tout au long de l'utilisation de l'adaptation par analogie : c'est l'approche *statique*. Dans un second temps, nous adopterons une approche *dynamique* de la définition de la fenêtre d'analogie, en cherchant à l'adapter de façon automatique aux caractéristiques de la situation cible.

Dans tout ce chapitre, du fait des résultats précédemment acquis, nous nous plaçons dans le contexte suivant :

- **variable d'analogie** : un ou plusieurs géopotentiels (la distance d'analogie est alors obtenue par une moyenne arithmétique sur les critères d'analogie calculés sur chacun des géopotentiels),
- **critère d'analogie** : score $S1$,
- **nombre d'analogues retenues** : 50 (méthode de référence).

5.1 Fenêtres d'analogie statiques

Dans toute cette section, nous allons nous intéresser aux fenêtres d'analogie statiques. Par fenêtre statique, nous entendons que la même fenêtre est utilisée pour toutes les situations cibles de la période de calibration. Elles sont donc indépendantes de la circulation atmosphérique propre à la situation cible considérée.

5.1.1 Représentativité de chaque point de grille

5.1.1.1 Cartographie de la pertinence des points de grilles

Avant d'entamer la recherche de la fenêtre la plus performante, il est intéressant de se pencher sur la représentativité de chacun des points de la grille sur laquelle sont définis nos champs. En fonction de leur éloignement par rapport aux groupements pluviométriques, on peut en effet s'attendre à ce que tous n'apportent pas une information aussi pertinente. Or, certaines caractéristiques de la représentativité spatiale de la circulation synoptique peuvent rester masquées si l'on considère une fenêtre d'analogie qui agrège d'emblée un nombre important de points de grille.

C'est pourquoi nous allons tout d'abord considérer la fenêtre d'analogie la plus petite possible. Toutefois, étant donné que le critère $S1$ porte sur les gradients des variables d'analogie, nous ne pouvons pas utiliser une fenêtre d'analogie qui ne concernerait qu'un seul point de grille. Nous choisissons donc comme *fenêtre élémentaire* un domaine rectangulaire de 2×2 points de grille selon la longitude et la latitude (soit un domaine de $2.5 \times 2.5^\circ$ en utilisant l'archive LTRE, basée sur les réanalyses NCEP/NCAR).

En utilisant cette fenêtre élémentaire, placée de façon arbitraire, comme fenêtre d'analogie, nous sommes en mesure de produire une série de prévisions dont nous pouvons ensuite évaluer l'utilité. Celle-ci reflète la pertinence des 4 points de grille considérés. Cette opération est en-

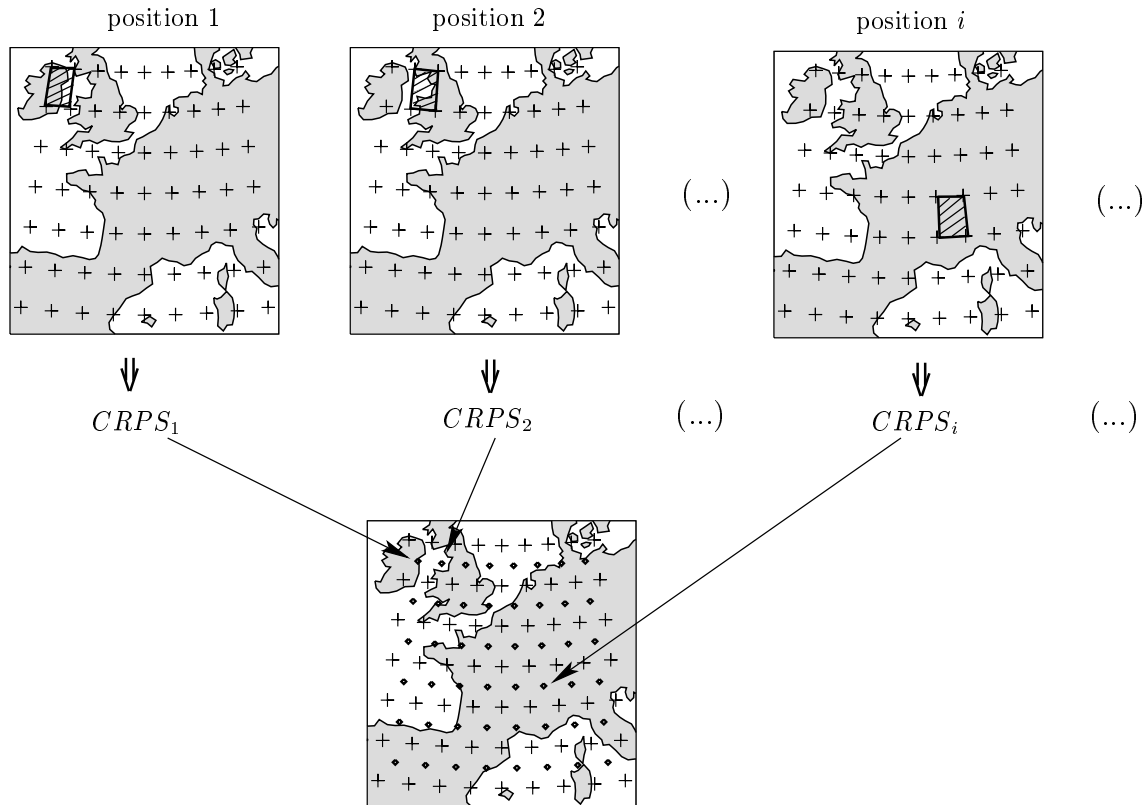


FIG. 5.1: Schéma de principe de la réalisation d'une cartographie de la pertinence de l'analogie.

suite répétée en déplaçant successivement la fenêtre élémentaire, de façon à parcourir l'intégralité du domaine d'archivage. Chacune des utilités ainsi obtenues sur la période choisie, peut finalement être associée au barycentre de la fenêtre utilisée ; nous construisons de cette manière une cartographie de la pertinence de l'analogie (figure 5.1).

5.1.1.2 Utilisation d'un seul géopotentiel

Cette méthode d'analyse par cartographie de la pertinence a été appliquée en considérant une variable d'analogie basée d'abord sur un seul géopotentiel. Celui-ci a été choisi de façon à mettre en évidence l'incidence du choix :

- du niveau, à heure de mesure fixée (+12h) : figure 5.3 a,
- de l'heure de mesure, à niveau fixé (850 hPa) : figure 5.3 b.

La première remarque que l'on peut faire au vu de ces résultats, obtenus sur l'ensemble de la période de calibration, est que le domaine sur lequel l'analogie est pertinente est assez restreint. Lorsque l'on considère une fenêtre d'analogie élémentaire suffisamment éloignée de nos groupements, la compétence du *CRPS* tend vers zéro : la circulation atmosphérique alors prise en compte n'est plus reliée aux précipitations que l'on mesure sur le groupement. Dans ces conditions, sélectionner des analogues revient à piocher au hasard dans la distribution climatologique. On retrouve ici le concept d'analogie à aire limitée mis en avant par Van Den Dool (1989). L'information ne se propage en effet pas instantanément. Ces cartographies montrent où se situe, à l'heure de mesure du géopotentiel d'analogie, l'information qui va concerner nos groupements entre +6h et +30h. Si la période de cumul des précipitations avait été inférieure à 24 heures, on peut supposer que les points de grille pertinents auraient été encore plus resserrés autour des groupements.

Si l'on considère le géopotentiel $z_{1000}+12$, on peut estimer que l'on rencontre de l'information pertinente jusqu'à une distance de 25 à 30° de longitude par rapport à la position des groupe-

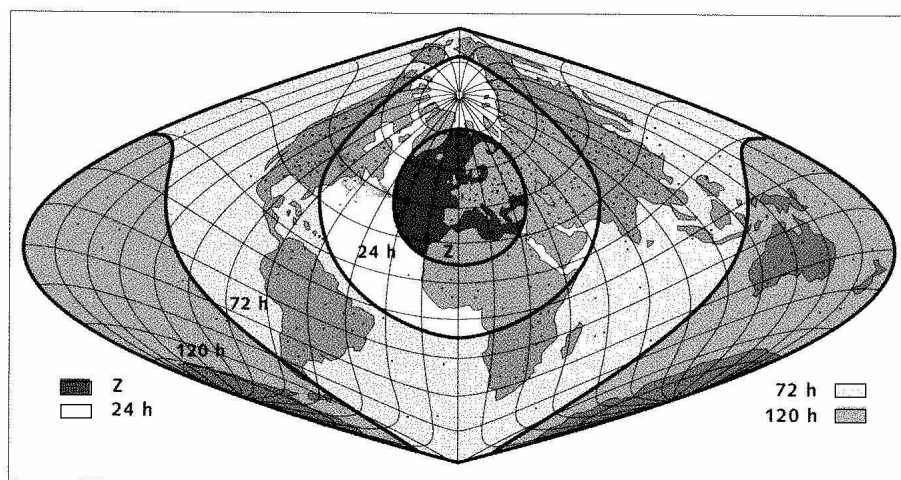


FIG. 5.2: Répartition des régions sur lesquelles des observations sont nécessaires pour réaliser des prévisions à échéance de 1, 3 et 5 jours sur la zone centrale Z (document CEPMMT, cité par Coiffier 2000).

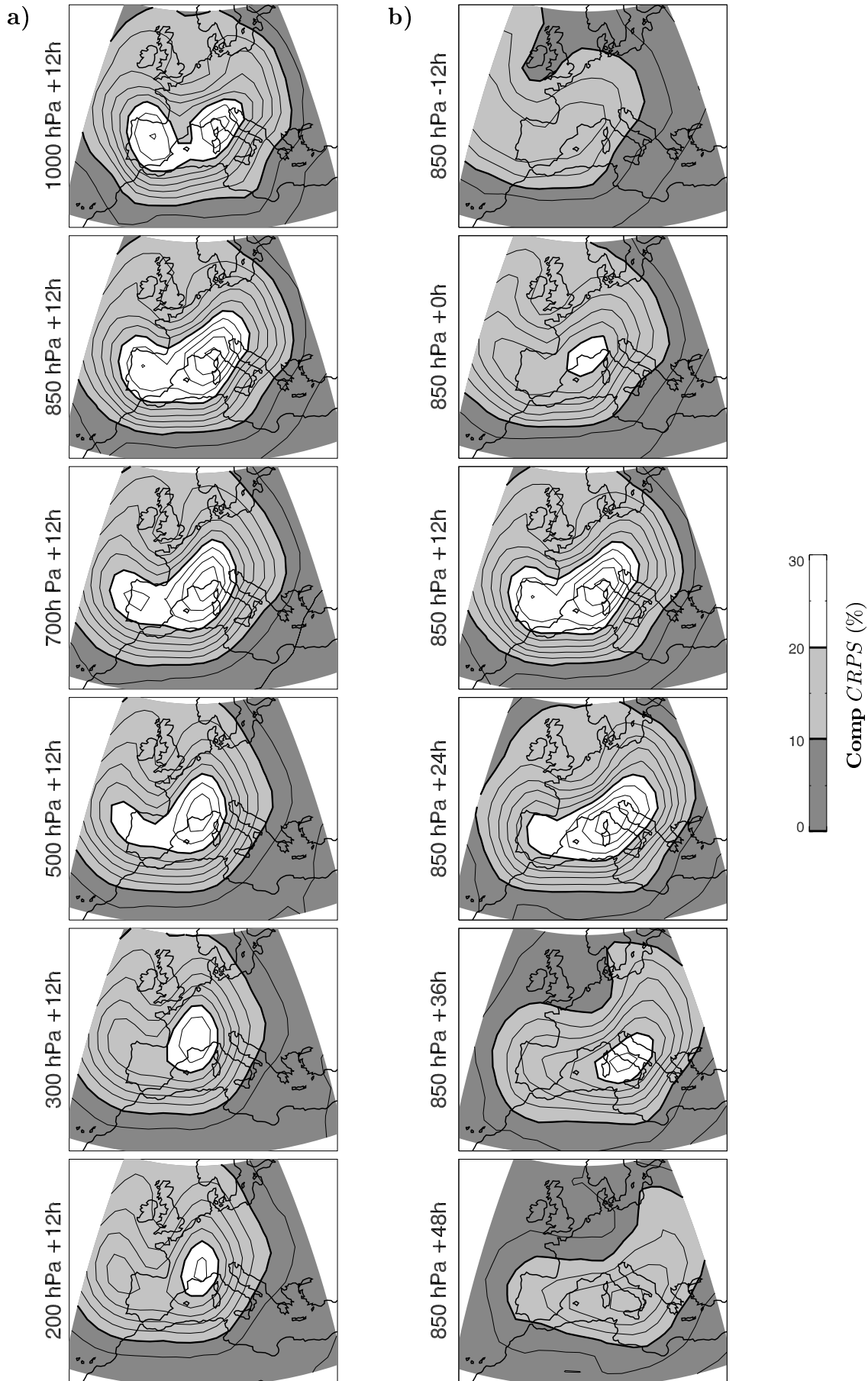


FIG. 5.3: Cartographie des performances (Compétence du *CRPS*) obtenues sur l'ensemble des groupements en fonction de la position d'une fenêtre d'analogie de $2.5 \times 2.5^\circ$. Variable d'analogie : un géopotentiel **a)** mesuré à +12h et dont le niveau varie de 1000 à 200 hPa, **b)** situé à 850 hPa et dont l'heure de mesure varie de -12h à +48h.

Note : les zones les plus claires correspondent aux zones les plus pertinentes.

ments. Il est intéressant de comparer ces résultats à ceux que l'on obtient avec un modèle NWP : cette distance est alors de 45° de longitude pour réaliser une prévision à échéance de 24 heures (figure 5.2). Il n'y a cependant pas de contradiction entre ces chiffres. Notre plage de prévision par rapport à z1000+12 s'étend en effet de -6 à +18h, si bien qu'en ce qui concerne l'adaptation par analogie :

- d'une part l'échéance est moins lointaine,
- et d'autre part le prédictand n'est pas un état à +24h, mais le cumul de tous les états entre -6 et +18h. L'exactitude de l'état à +18h ayant une incidence faible sur le cumul global, les points de grille les plus éloignés, pertinents pour prévoir ces échéances, apportent finalement peu d'information.

On remarque ensuite que la répartition de la pertinence des points de grille n'est pas isotrope. Le domaine le plus informatif décrit une sorte de U ou de L « *couché sur le dos* » selon le géopotentiel considéré, avec une zone peu significative centrée sur les îles britanniques. Cette répartition semble faire écho à celle des groupements pluviométriques EDF, lesquels ne sont pas non plus répartis de façon homogène et se regroupent principalement le long de la vallée du Rhône et sur les Pyrénées.

Le choix du niveau du champ de hauteurs géopotentielles considéré a très peu d'influence sur les configurations obtenues, mis à part un changement global de performance moyenne et un très léger décalage vers l'Ouest avec l'augmentation de l'altitude du niveau de mesure.

L'heure de mesure exerce par contre une influence certaine sur l'information que peut apporter chaque point de grille. La figure 5.3b montre clairement le glissement vers l'Est du domaine pertinent avec l'augmentation de l'heure de mesure. Comme nous l'avons déjà évoqué précédemment (cf. 4.2.2.1), cela peut aisément s'interpréter grâce à une description lagrangienne du déplacement moyen des masses d'air sous nos latitudes. Elles se déplacent en effet généralement d'Ouest en Est. Ainsi, une masse d'air située à l'aplomb de nos groupements à +12h se trouvait souvent plus à l'Ouest à +0h ou -12h, alors qu'elle déplacera vers l'Est pour les heures de mesure ultérieures. Il apparaît donc comme important que l'analogie porte sur les masses d'air qui concerneront (ou concernaient, selon l'heure de mesure choisie) les groupements lors de la période de cumul des précipitations.

Les choix de la fenêtre d'optimisation et de l'heure de mesure du géopotentiel sont donc liés entre eux. Si l'on choisit une heure de mesure qui ne se trouve pas dans la période de cumul, alors l'analogie doit porter sur une zone excentrée par rapport aux groupements. Cependant, le déplacement des masses d'air de l'Ouest vers l'Est n'est qu'une moyenne, qui a donc vocation à être prise en défaut pour chaque cas particulier. Cette analogie sur une fenêtre excentrée est donc globalement moins performante qu'une analogie effectuée à +12h ou +24h à l'aplomb des groupements.

5.1.1.3 Combinaison de plusieurs géopotentiels

Lors du chapitre précédent, il est apparu que la combinaison de plusieurs géopotentiels pouvait apporter un gain significatif de performance. Nous avons donc repris la cartographie de

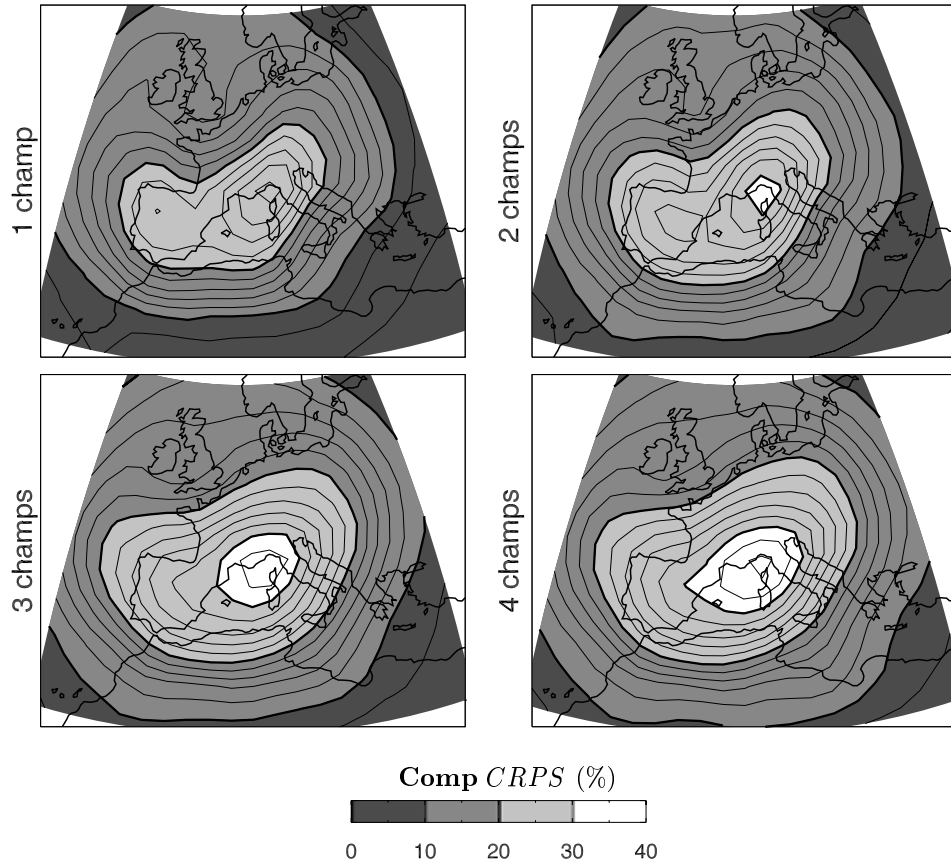


FIG. 5.4: Cartographie de la pertinence de l'analogie pour les combinaisons de géopotentiels les plus performantes : **a)** : 1 champ : $z_{850}+12$; **b)** : 2 champs : $z_{1000}+12$ et $z_{500}+24$; **c)** : 3 champs : $z_{1000}+12$, $z_{500}+12$ et $z_{700}+36$; **d)** : 4 champs : $z_{1000}+12$, $z_{500}+12$, $z_{1000}+24$ et $z_{500}+36$.

la pertinence de l'analogie, en considérant successivement les meilleures combinaisons obtenues pour 1, 2, 3 et 4 champs (figure 5.4).

A part en ce qui concerne le niveau moyen de performance, on constate assez peu de variations dans les patterns obtenus en fonction du nombre de géopotentiels servant de support à l'analogie. La principale conséquence de son augmentation réside dans un « lissage » des structures spatiales obtenues : la répartition de la pertinence est plus isotrope autour des groupements considérés.

5.1.2 Algorithme d'optimisation

Après avoir étudié la pertinence des points de grille considérés individuellement, nous allons agréger ceux-ci afin d'identifier la fenêtre d'analogie qui contient le plus d'information. Toutefois, l'ensemble des possibilités existantes est trop important pour que nous puissions les tester toutes exhaustivement. Il nous faut donc faire des choix et nous restreindre à certains types de fenêtre.

Les résultats précédents ont montré que la répartition de la pertinence, bien que non isotrope, est assez lisse, surtout lorsque l'on considère plus d'un géopotential comme variable d'analogie. C'est pourquoi nous décidons de nous limiter à des fenêtres rectangulaires. Ce parti pris nous permet de définir une fenêtre selon 4 paramètres (figure 5.5) :

- **lon0** : longitude du coin inférieur gauche,
- **lat0** : latitude du coin inférieur gauche,
- **dlon** : longueur, en degré de longitude, du côté zonal (longueur Est-Ouest),
- **dlat** : longueur, en degré de latitude, du côté méridien (longueur Nord-Sud).

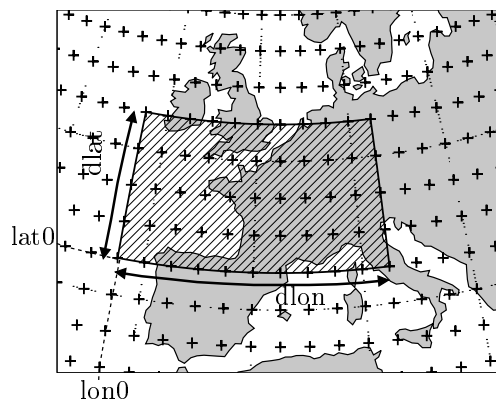


FIG. 5.5: Représentation d'une fenêtre d'analogie en fonction des paramètres lon0 , lat0 , dlon et dlat . Dans cet exemple, $\text{lon0}=12.5^\circ\text{W}$, $\text{lat0}=42.5^\circ\text{N}$, $\text{dlon}=25^\circ$ et $\text{dlat}=10^\circ$.

Afin d'identifier la fenêtre la plus pertinente, nous partons d'une fenêtre de petite taille que nous faisons grandir par itérations successives, de manière à ce que sa croissance se fasse à chaque étape dans la direction qui apporte le plus d'information :

- ① Nous considérons tout d'abord la fenêtre élémentaire la mieux centrée par rapport au groupement (ou ensemble de groupements) utilisé(s) : elle forme notre première *fenêtre étape*.
- ② Pour chacune des 4 directions (Nord, Sud, Est et Ouest), nous créons une nouvelle fenêtre en agrandissant la fenêtre étape d'un pas de grille : ce sont *les fenêtres agrandies*.
- ③ Chacune des fenêtres agrandies est alors utilisée successivement comme fenêtre d'analogie pour « *rejouer* » la prévision sur l'ensemble de la période de calibration. On obtient ainsi un score d'utilité pour chacune d'elles.
- ④ La fenêtre agrandie qui présente les meilleures performances devient notre nouvelle fenêtre étape, avec laquelle nous reprenons l'étape ②.

Les étapes ② et ③ sont ensuite répétées jusqu'à atteindre les limites du domaine d'archivage. Finalement, la fenêtre étape qui s'est révélée la plus pertinente constitue notre fenêtre d'analogie optimale. On peut considérer par ailleurs que chacune des fenêtres étapes constitue un optimum pour une taille de fenêtre donnée.

Ce type d'algorithme, qui fait évoluer les paramètres d'un modèle selon un chemin unique, peut conduire à la meilleure solution à condition qu'il n'y ait pas d'optima locaux. Nous pouvons raisonnablement penser que c'est bien le cas ici étant donné la régularité des isolignes affichées par les cartographies de pertinence. Il convient néanmoins de noter que le choix de la fenêtre initiale n'est pas anecdotique. Les points constitutifs de celle-ci feront en effet nécessairement partie de la fenêtre retenue comme optimale.

5.1.3 Fenêtres communes à l'ensemble des groupements

Dans toute cette sous-section, nous allons appliquer l'algorithme de recherche de fenêtres optimales en utilisant la même fenêtre d'analogie pour l'ensemble des groupements pluviométriques EDF. Ce sont donc les mêmes situations qui seront considérées comme analogues à une situation cible donnée, que le groupement qui nous intéresse soit situé en Bretagne ou en Corse.

5.1.3.1 Fenêtres uniques

Dans un premier temps, pour une combinaison de géopotentiels donnée, nous recherchons une fenêtre optimale unique, commune à l'ensemble des champs pris en compte. Nous sommes donc dans le même contexte que lors du chapitre précédent, lorsque nous nous intéressions au choix des géopotentiels.

Nous considérons ainsi les fenêtres d'analogie optimales associées aux meilleures combinaisons de géopotentiels obtenues pour 1, 2, 3 et 4 champs, dont les paramètres sont résumés dans le tableau 5.1. La figure 5.6 présente en outre, pour chacune de ces combinaisons, l'évolution des performances obtenues en fonction de la longueur diagonale des différentes fenêtres étapes considérées successivement par l'algorithme d'optimisation.

On confirme l'impression donnée par les cartographies de pertinence, à savoir que les fenêtres optimales sont assez restreintes, de l'ordre de $25^\circ \times 10^\circ$. Il n'est même pas nécessaire de prendre en compte tous les points apportant de l'information. Seuls les plus pertinents, et donc les plus proches des groupements, sont à considérer. Le graphe 5.6 montre cependant que les performances diminuent lentement si l'on considère une fenêtre légèrement trop grande, alors que l'utilisation d'une fenêtre trop petite pénalise fortement les résultats. En effet, si l'on agrandit la fenêtre, on ne fait que diluer l'information réellement pertinente par de l'information moins pertinente ou même sans influence sur le prédictand. Au contraire, si on la diminue, on risque d'ignorer une partie de la circulation synoptique qui intéresse directement les groupements pluviométriques.

Il convient de noter que ces fenêtres ne sont pas centrées sur le centre de gravité de nos groupements. Elles sont toutes décalées vers l'Ouest et le Sud. Or, ce sont de ces directions que proviennent principalement les flux générateurs de précipitations intéressant nos groupements. Ainsi, ceux situés dans les Alpes du Nord, la zone jurassienne, le versant Est du massif central et la partie occidentale des Pyrénées sont surtout soumis à des précipitations d'origine océanique, alors que ceux des Alpes du Sud, des Cévennes et du Vivarais sont majoritairement exposés

Géopotentiels utilisés	Paramètres de la fenêtre optimale				Performance
	lon0 (°E)	lat0 (°N)	dlon (°)	dlat (°)	Comp CRPS (%)
z850+12	-10	37.5	27.5	10	35.33
z1000+12, z500+24	-7.5	37.5	25	10	38.23
z1000+12, z500+12, z700+36	-5	37.5	20	10	38.85
z1000+12, z500+12, z1000+24, z500+36	-5	37.5	22.5	10	39.45

TAB. 5.1: Fenêtres d'analogie optimales pour les meilleures combinaisons de 1, 2, 3 et 4 géopotentiels.

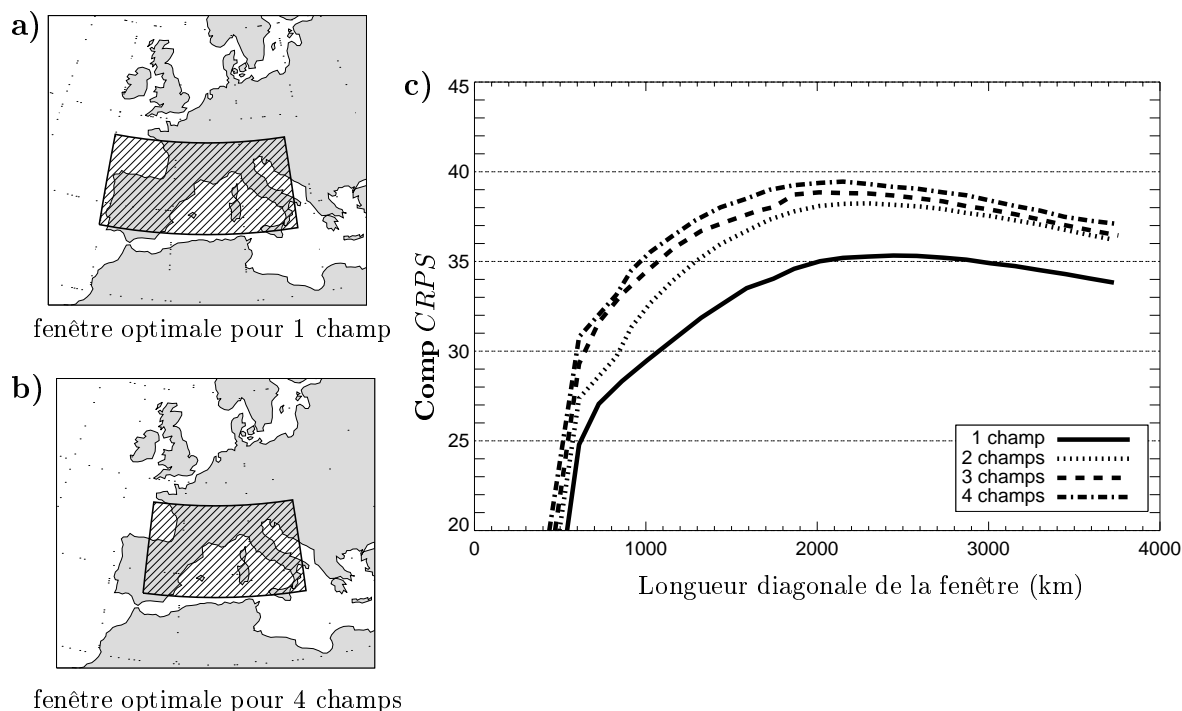


FIG. 5.6: Recherche d'une fenêtre d'analogie unique optimale. **a)** : fenêtre optimale pour $z800+12$; **b)** : fenêtre optimale pour $(z1000+12, z500+12, z1000+24, z500+36)$; **c)** : évolution des performances obtenues en fonction de la longueur diagonale des différentes fenêtres étapes, pour chacune des combinaisons de géopotentiels considérées.

à des flux de Sud qui proviennent de la Méditerranée, du moins en ce qui concerne les plus fortes précipitations. L'analogie se focalise donc sur ce qui « arrive » sur le groupement d'intérêt. Notons néanmoins que les groupements situés sur le versant italien des Alpes sont pour leur part surtout influencés par des systèmes provenant du Sud-Est, remontant la vallée du Pô. Or, ceci n'apparaît pas dans les grilles optimales, peut-être parce que ces groupements sont trop isolés, et donc peu représentatifs du comportement moyen de l'ensemble.

Le décalage progressif de la fenêtre optimale vers l'Est, que l'on observe avec l'augmentation du nombre de géopotentiels servant de support à l'analogie, confirme nos hypothèses quant à l'apparition de champs à $+36h$ dans les combinaisons optimales. Ces champs servaient à corriger le caractère excentré de la fenêtre alors utilisée. Leur heure de mesure tardive permet de décaler artificiellement la fenêtre vers l'Est quand on est, comme ici, dans un contexte de flux dominant venant de l'Ouest. Ce décalage s'accompagne par ailleurs d'une diminution de la taille de la fenêtre optimale : le fait de considérer simultanément plusieurs heures de mesure permet de considérer des domaines plus limités. En effet, ce qui se trouve à l'Ouest loin des groupements à $+12h$, s'en est rapproché à $+24h$.

Finalement, aussi bien les performances atteintes que la fenêtre optimale semblent converger quand on augmente le nombre de géopotentiels. Il semble bien que l'on atteigne une limite et que l'augmentation de performance liée à l'utilisation de plus que deux champs ne soit pas réellement significative.

5.1.3.2 Fenêtre propre à chaque champ

Dans un second temps, nous nous autorisons à appliquer une fenêtre d'analogie différente pour chacun des géopotentiels considérés. Ainsi, si la variable d'analogie est une combinaison de trois géopotentiels, alors nous manipulons trois fenêtres différentes : une pour chaque champ.

Nous employons le même algorithme d'optimisation que précédemment, en faisant varier ensemble les paramètres de toutes les fenêtres utilisées. Nous avons ainsi plusieurs fenêtres étapes en même temps. A chaque itération, seule celle qui possède la fenêtre agrandie la plus performante se voit modifiée.

Les résultats des optimisations, appliquées ici à nouveau aux meilleures combinaisons de 1, 2, 3 ou 4 géopotentiels, sont présentés par la figure 5.8 et comparés avec ceux obtenus lors de l'utilisation de fenêtres uniques, par la figure 5.7.

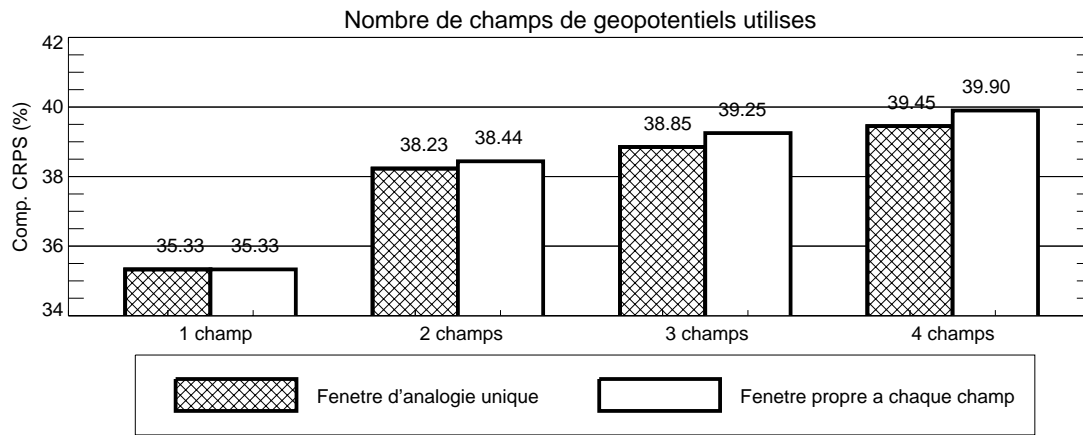


FIG. 5.7: Comparaison des performances obtenues en utilisant une fenêtre d'analogie unique ou des fenêtres propres à chaque champ, pour chacune des meilleures combinaisons de 1, 2, 3 ou 4 géopotentiels.

Les fenêtres optimales obtenues pour la combinaison de deux champs diffèrent peu entre elles. Celle associée à $z1000+12$ correspond quasiment à celle obtenue avec un seul géopotentiel, alors que celle associée à $z500+24$ semble plus orientée Nord-Sud. Les dimensions de cette dernière ne sont toutefois pas réellement significatives, étant donné que passé un certain seuil (longueur diagonale ≈ 1500 km), son agrandissement est quasiment sans influence sur les performances.

Au delà de deux champs, l'interprétation des fenêtres optimales obtenues pose un réel problème. Elles apparaissent comme plus ou moins aléatoires et ne semblent pas revêtir la moindre signification physique. Certaines ne couvrent qu'une étroite bande de latitude, alors que d'autres atteignent des dimensions exagérées. L'algorithme d'optimisation peine en outre à trouver une direction optimale : il n'existe pas une combinaison unique de fenêtre optimale. Ce constat confirme le caractère artificiel du faible gain de performance obtenu lorsque l'on considère plus de deux géopotentiels.

La différenciation des fenêtres d'analogie en fonction des géopotentiels utilisés ne semble donc absolument pas pertinente, pas plus que ne le sont les gains de performance associés.

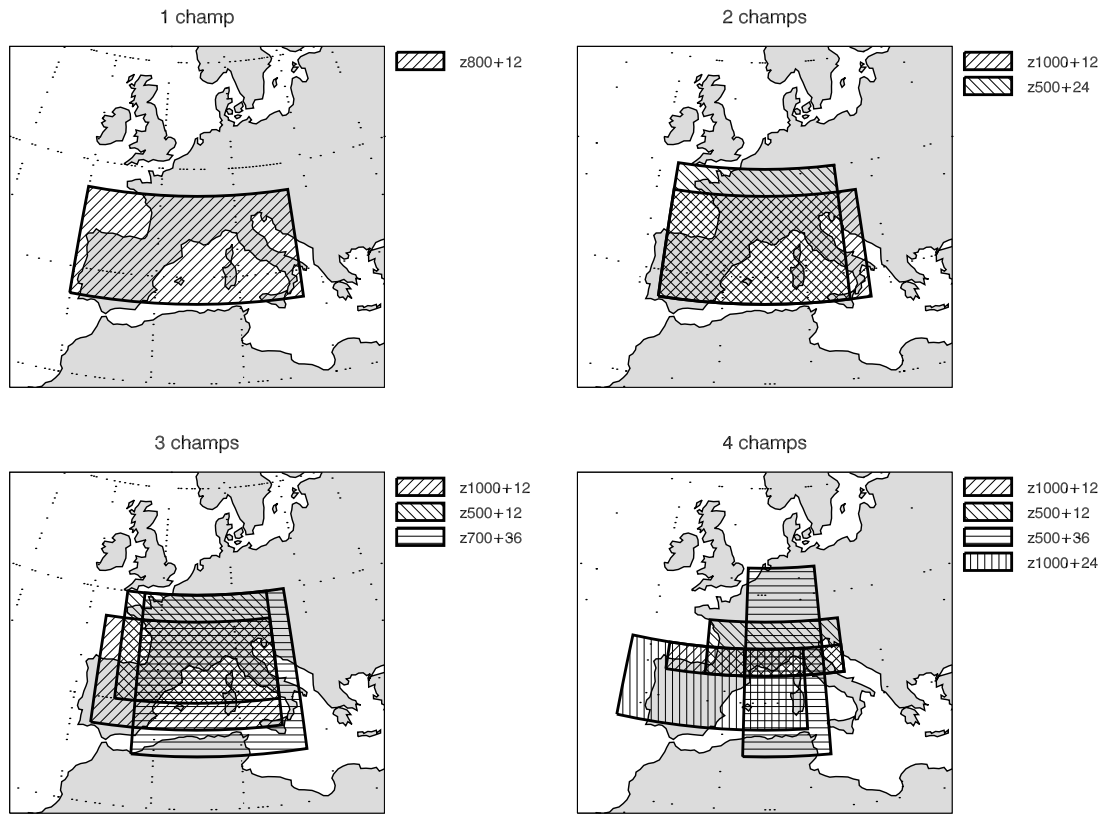


FIG. 5.8: Recherche de fenêtres d'analogie propres à chaque champ, pour chacune des meilleures combinaisons de 1, 2, 3 ou 4 géopotentiels.

Finalement, cette approche alourdit notablement les algorithmes de recherche d'analogues pour un gain négligeable par rapport à l'utilisation de fenêtres uniques.

5.1.4 Dépendance à la position géographique du groupement

L'étude, lors du chapitre précédent (cf. 4.2.2.2), de la répartition spatiale des performances obtenues sur chaque groupement, a montré entre autres choses, que plus on s'éloignait du centre de gravité de l'ensemble des groupements, moins les performances étaient élevées. Une fenêtre d'analogie commune à tous les groupements devient en effet excentrée par rapport à ceux situés en marge des autres. Ainsi, les fenêtres uniques optimales calculées précédemment semblent complètement ignorer le groupement **Bretagne**, qui se trouve fortement isolé.

Il peut donc être intéressant de considérer plusieurs fenêtres d'analogie en fonction du groupement sur lequel porte la prévision. Mais dans ce cas, est-il forcément indispensable de considérer une fenêtre différente pour chaque groupement, ce qui multiplie par autant les temps de calcul, ou est-il possible de *régionaliser* les groupements afin de limiter le nombre de fenêtres utilisées ?

Nous allons tenter de répondre à cette question, en considérant comme variable d'analogie le couple de géopotentiels ($z1000+12$, $z500+24$).

5.1.4.1 A chaque groupement, sa fenêtre !

Dans un premier temps, nous prenons le luxe de calculer quelle est la fenêtre d'analogie optimale pour chaque groupement pluviométrique. A partir des 52 groupements EDF, cela nous amène à identifier pas moins de 28 fenêtres optimales différentes.

Si l'on compare les performances obtenues en utilisant ces 28 fenêtres individuelles avec celles obtenues en se contentant d'une fenêtre commune à tous les groupements (figure 5.9), il apparaît que pour bon nombre de groupements, la différence est bien maigre. Or, les fenêtres individuelles font généralement de l'ordre de $20^\circ \times 10^\circ$, c'est-à-dire qu'elles sont de taille légèrement inférieure à celle de la fenêtre commune. Donc, pour tous les groupements situés à proximité du centre de gravité, pour lesquels la fenêtre commune est bien centrée, l'utilisation de cette dernière revient seulement à considérer un domaine un peu trop grand par rapport à leur domaine optimal. Comme nous l'avons vu précédemment, cela pénalise peu les performances.

Seuls les groupements situés plus en marge, tels que les groupements bretons, pyrénéens ou italiens, montrent une nette amélioration de leurs performances. En ce qui les concerne, les influences climatiques qui les intéressent sont trop éloignées des influences majoritaires sur l'ensemble des groupements pour que la fenêtre commune soit proche d'une fenêtre optimale.

Ainsi, même si les fenêtres individuelles ne révolutionnent pas les prévisions lorsque celles-ci sont considérées dans leur ensemble (la compétence du *CRPS* passe de 38.23% à 39.29%), elles évitent que certains groupements ne présentent d'importants déficits de performance (la compétence du *CRPS* passe ainsi sur **Bretagne** de 32.74% à 36.24%).

5.1.4.2 Régionalisation des groupements

L'utilisation de fenêtres individuelles demande une débauche de moyen qui n'est pas forcément raisonnable... Or, les résultats précédents montrent que seuls les groupements les plus éloignés souffrent réellement de l'utilisation d'une fenêtre commune. On peut donc espérer qu'en regroupant entre eux des bassins proches les uns des autres, pour lesquels nous utiliserions une même fenêtre d'analogie, nous réussirons à la fois à nous restreindre à un nombre raisonnable de fenêtres et à obtenir des performances peu dégradées.

Nous cherchons donc à procéder à une régionalisation de nos groupements. Mais pour cela, jusqu'à quelle distance est-il donc raisonnable d'utiliser une fenêtre commune pour rechercher des analogues sur 2 groupements distincts ?

Pour essayer d'apporter des éléments de réponse, nous avons considéré l'ensemble des 1326 couples possibles en considérant 2 à 2 les 52 groupements EDF. Pour chacun de ces couples, nous avons identifié la fenêtre commune optimale. Il nous est donc possible de calculer la perte relative de performances sur chaque couple ΔS , consécutive à l'utilisation de la fenêtre commune f_com plutôt qu'à celle des fenêtres individuelles f_ind_i de chacun des groupements i du couple :

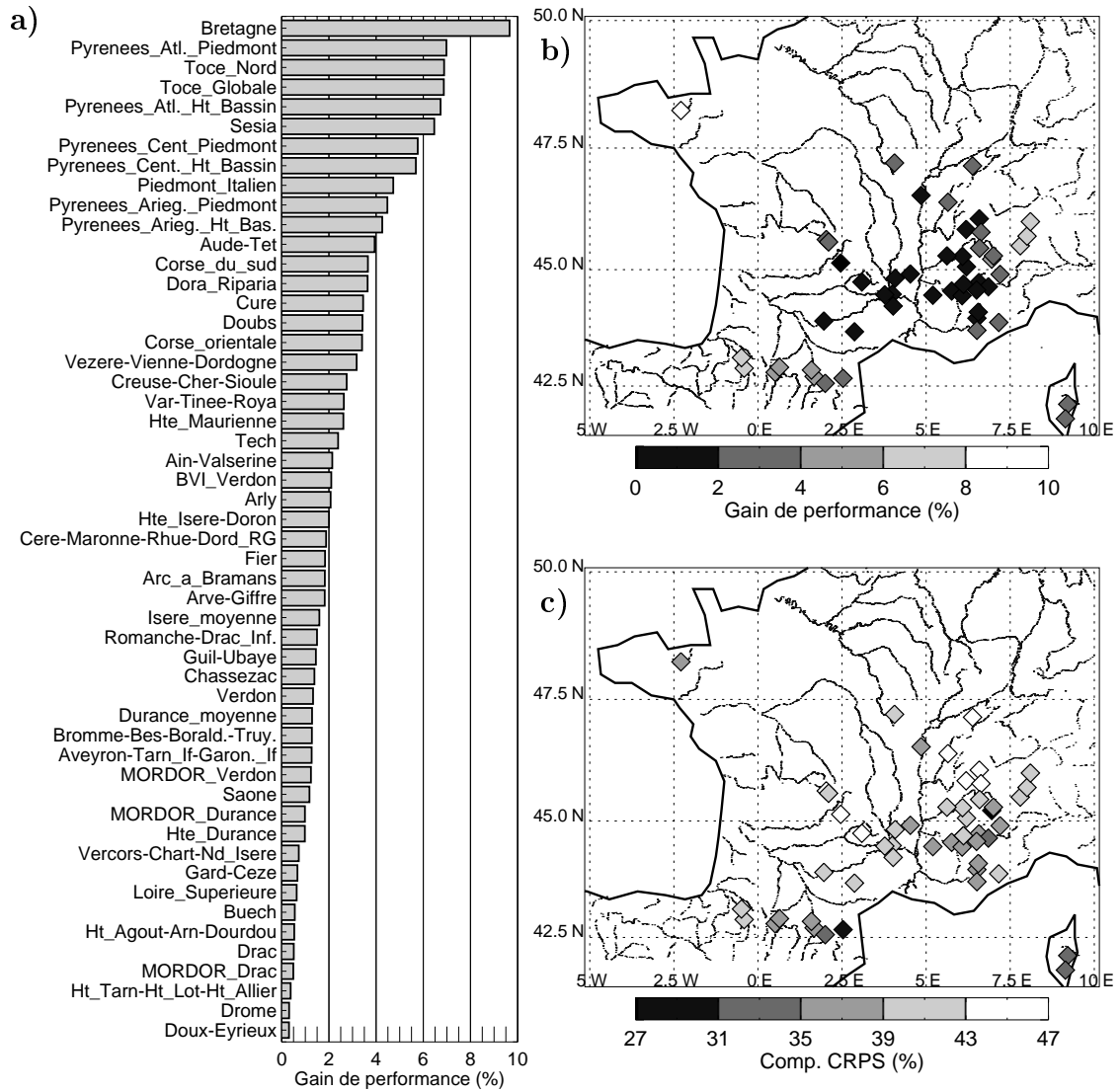


FIG. 5.9: Comparaison des performances obtenues en utilisant les fenêtres optimales individuelles ou une fenêtre commune. **a)** et **b)** : Gain de performance en fonction de chaque groupement (exprimé en pourcentage de **Comp CRPS**) ; **c)** : performances atteintes en utilisant les fenêtres individuelles.

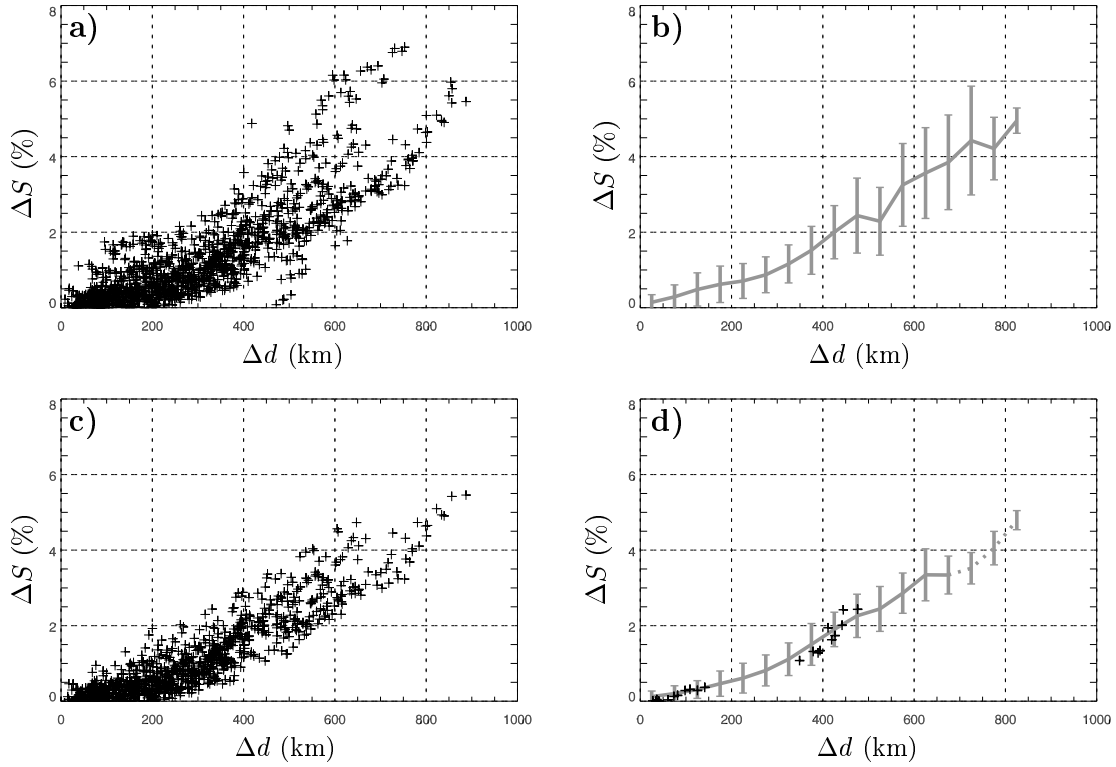


FIG. 5.10: Perte de performance ΔS enregistrée sur un couple de groupements en utilisant une fenêtre commune à la place de fenêtres individuelles, en fonction de la distance Δd qui sépare ces groupements. **a)** : nuage de points, tous les couples ; **b)** : moyenne et écart-type par classe de distance, tous les couples ; **c)** : nuage de points, couples « hexagonaux » (cf. texte) ; **d)** : moyenne et écart-type, couples « hexagonaux » + nuage de points, couples « italiens ».

$$\Delta S = \frac{\frac{1}{2} \cdot (S_1^{f-ind_1} + S_2^{f-ind_2}) - \frac{1}{2} \cdot (S_1^{f-com} + S_2^{f-com})}{\frac{1}{2} \cdot (S_1^{f-ind_1} + S_2^{f-ind_2})} \quad (5.1)$$

où :

- $S_1^{f-ind_1}$ est la performance (exprimée en **Comp CRPS**) obtenue sur le premier groupement du couple, en utilisant sa fenêtre individuelle,
- $S_2^{f-ind_2}$ est la performance obtenue sur le second groupement du couple, en utilisant sa fenêtre individuelle,
- S_1^{f-com} est la performance obtenue sur le premier groupement du couple, en utilisant la fenêtre commune de ce couple,
- S_2^{f-com} est la performance obtenue sur le second groupement du couple, en utilisant la fenêtre commune de ce couple.

En reportant ces pertes de performances ΔS en fonction de la distance Δd qui sépare géographiquement les membres de chaque couple, on trace une sorte de variogramme empirique (figure 5.10).

Considéré sans distinction, l'ensemble de tous les couples forme un nuage de points qui présente une dispersion très importante (figure 5.10a et b). Dans le cas général, la perte de performance n'est donc pas liée à la seule distance qui sépare les groupements.

Néanmoins, si l'on soustrait de l'échantillon tous les couples qui font intervenir au moins un groupement situé :

- soit sur le versant italien des Alpes,
- soit en Corse,

pour ne conserver que les couples que nous appelons « *hexagonaux* », on voit apparaître un comportement beaucoup plus intéressant (figure 5.10c). Les couples hexagonaux affichent en effet une perte de performance fortement liée à la distance inter-groupements. De plus, si on considère maintenant uniquement les couples qui ne font intervenir que des groupements italiens ou corses (couples « *italiens* »), on s'aperçoit qu'ils suivent le même comportement que les couples hexagonaux (figure 5.10d). Il est à noter que les distances supérieures à 650 km ne sont pas significatives. Ces distances correspondent en effet systématiquement à des couples faisant intervenir le groupement **Bretagne**. L'échantillonnage est donc biaisé, ces différents couples n'étant pas indépendants.

Les Alpes se comportent donc comme une barrière climatique. Les situations présentes sur l'un des versants influent peu sur l'autre. Deux groupements situés sur deux versant opposés, même très proches l'un de l'autre, tels que les groupements **Hte_Isère_Doron** et **Piedmont_italien**, voient donc des régimes générateurs de précipitations dominants très différents : dans le premier cas, ils viennent de l'Ouest, alors que dans le second ils remontent la vallée du Pô. Il y a fort à parier que les Pyrénées constituent une barrière climatique semblable, mais nous n'avons malheureusement pas de groupement situé sur leur versant Sud pour le vérifier.

Ainsi, indépendamment de chaque côté des Alpes, la seule distance qui sépare les groupements permet d'estimer s'ils peuvent se contenter de la même fenêtre d'analogie. La présence de reliefs comme les Cévennes par exemple, ne semble pas avoir d'influence sur le choix de la fenêtre : seule compte la distance entre les groupements. La vallée du Rhône, qui crée une direction Nord-Sud privilégiée au sein de nos groupements, nous a incité à vérifier si le comportement des couples était identique selon qu'ils étaient disposés sur un axe Est-Ouest ou Nord-Sud (avec une tolérance de $\pm 10^\circ$). Les résultats (figure 5.11) ne s'écartent cependant pas significativement du comportement général des couples hexagonaux. Encore une fois, seule semble compter la distance entre les groupements.

On remarque que le variogramme obtenu présente une rupture de pente aux alentours de $\Delta d = 250$ km. Or, la grille sur laquelle sont définis les champs météorologiques de l'archive LTHE est de $2.5 \times 2.5^\circ$ (résolution des réanalyses NCEP/NCAR), c'est à dire qu'à nos latitudes, une maille représente à peu près 250 km. Les couples de groupements dont l'inter-distance est inférieure à 250 km ne présentent donc pas forcément des fenêtres individuelles différentes.

Finalement, la figure 5.10d nous permet de répondre à notre question. Si l'on se fixe un seuil de perte de performance tolérable ΔS , on peut estimer la distance à ne pas dépasser entre les groupements que l'on désire regrouper ensemble sous une même fenêtre d'analogie (à condition

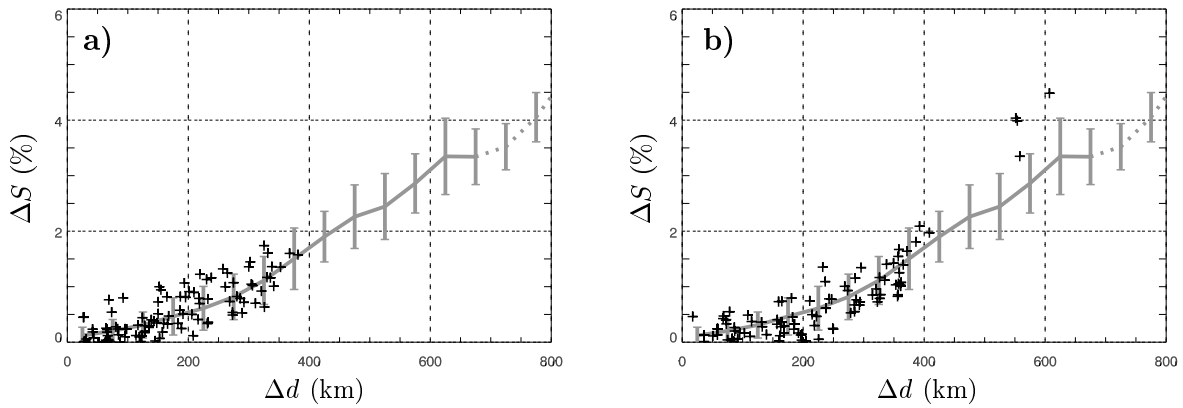


FIG. 5.11: Perte de performance ΔS enregistrée sur un couple de groupements en utilisant une fenêtre commune à la place de fenêtres individuelles, en fonction de la distance Δd qui sépare ces groupements. **a)** : moyenne et écart-type, couples hexagonaux + nuage de points des couples hexagonaux ayant une direction Nord-Sud; **b)** : moyenne et écart-type, couples hexagonaux + nuage de points des couples hexagonaux ayant une direction Est-Ouest.

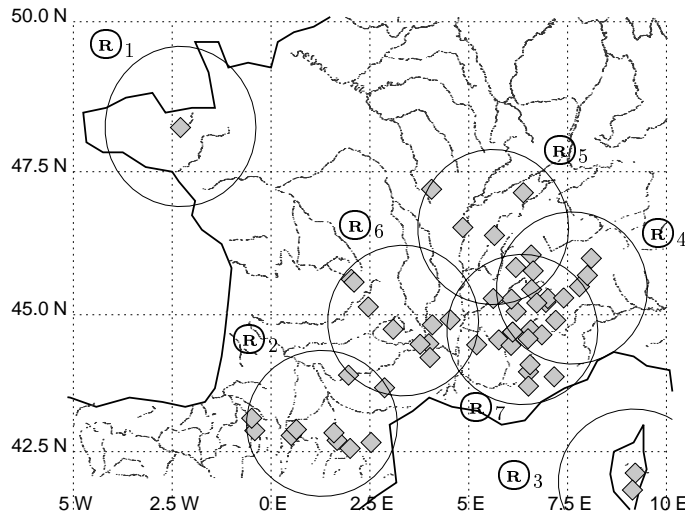


FIG. 5.12: Regroupement selon un diamètre maximal $\Delta d = 300$ km.

de ne pas franchir une barrière climatique). Par exemple, si l'on considère $\Delta S = 1.5\%$, alors on obtient $\Delta d \approx 300$ km : chaque regroupement doit tenir dans un cercle de 300 km de diamètre. Dans le cas présent, cet exemple nous amène à considérer 7 fenêtres différentes (figure 5.12 et tableau 5.2).

En considérant ces 7 fenêtres, on obtient un score global **Comp CRPS** = 39.03%, contre 38.23% avec une fenêtre commune et 39.29% avec les 28 fenêtres individuelles. La plus grande perte de performance n'est par ailleurs plus que de 2.34% (groupement *Vezere-Vienne-Dordogne*) alors qu'elle était de 9.82% avec une fenêtre commune (groupement *Bretagne*). Ainsi, même si l'utilisation d'une fenêtre commune sur un tel ensemble de groupements est préjudiciable, il

Regroupement	Paramètres de la fenêtre optimale			
	lon0 (°E)	lat0 (°N)	dlon (°)	dlat (°)
\textcircled{R}_1	-15	40	20	12.5
\textcircled{R}_2	-10	35	22.5	12.5
\textcircled{R}_3	-2.5	35	20	10
\textcircled{R}_4	2.5	37.5	15	12.5
\textcircled{R}_5	-5	40	20	10
\textcircled{R}_6	-7.5	37.5	22.5	10
\textcircled{R}_7	-5	37.5	22.5	10

TAB. 5.2: Fenêtres d'analogie optimales pour les regroupements effectués selon un seuil $\Delta d = 300$ km.

n'est pas besoin de considérer pour autant des fenêtres individuelles. Un regroupement judicieux permet de trouver un compromis entre légèreté des algorithmes et perte de performances.

5.2 Conditionnement par la situation cible

Jusqu'à présent, nous n'avons considéré que des fenêtres d'analogie statiques. Toutefois, il est apparu que leur positionnement optimal n'était pas exactement à l'aplomb des groupements, mais plutôt décalé en direction de la provenance dominante des situations génératrices de précipitations. On peut donc légitimement se demander s'il ne serait pas avantageux d'adapter la fenêtre utilisée à la circulation météorologique propre à la situation cible.

5.2.1 Constat préliminaire

Cette question du conditionnement de la fenêtre par la situation cible nous est apparue lorsque nous nous sommes penchés sur les changements qui apparaissent dans les cartographies de pertinence, selon le mois de l'année que l'on considère. A titre d'exemple, la figure 5.13 présente les cartographies de pertinence obtenues mois par mois sur le groupement **Arc_à_Bramans**.

On distingue clairement une évolution de la pertinence en fonction de la période de l'année. Ainsi, lors des mois d'hiver, soumis principalement aux flux d'Ouest, les points de grille les plus pertinents se trouvent effectivement à l'Ouest de notre groupement. Le printemps et le début de l'automne, qui sont des périodes propices aux fortes pluies venant du Sud, affichent pour leur part des points de grilles pertinents alignés selon un axe Nord-Sud.

La période estivale se démarque du reste de l'année par une pertinence à la fois faible et peu organisée spatialement. C'est la période où l'adaptation par analogie sur la circulation synoptique est la moins performante. En effet, les pluies estivales sont principalement générées par des orages locaux. Or, ceux-ci sont peu contrôlés par la circulation synoptique : même si le risque orageux peut se reconnaître à grande échelle, leur occurrence ou non dépend principalement des conditions locales dans les très basses couches de l'atmosphère.

La première application de ces résultats pourrait être de chercher à optimiser une fenêtre différente pour chaque mois de l'année. Par exemple, si la situation cible est un 3 octobre,

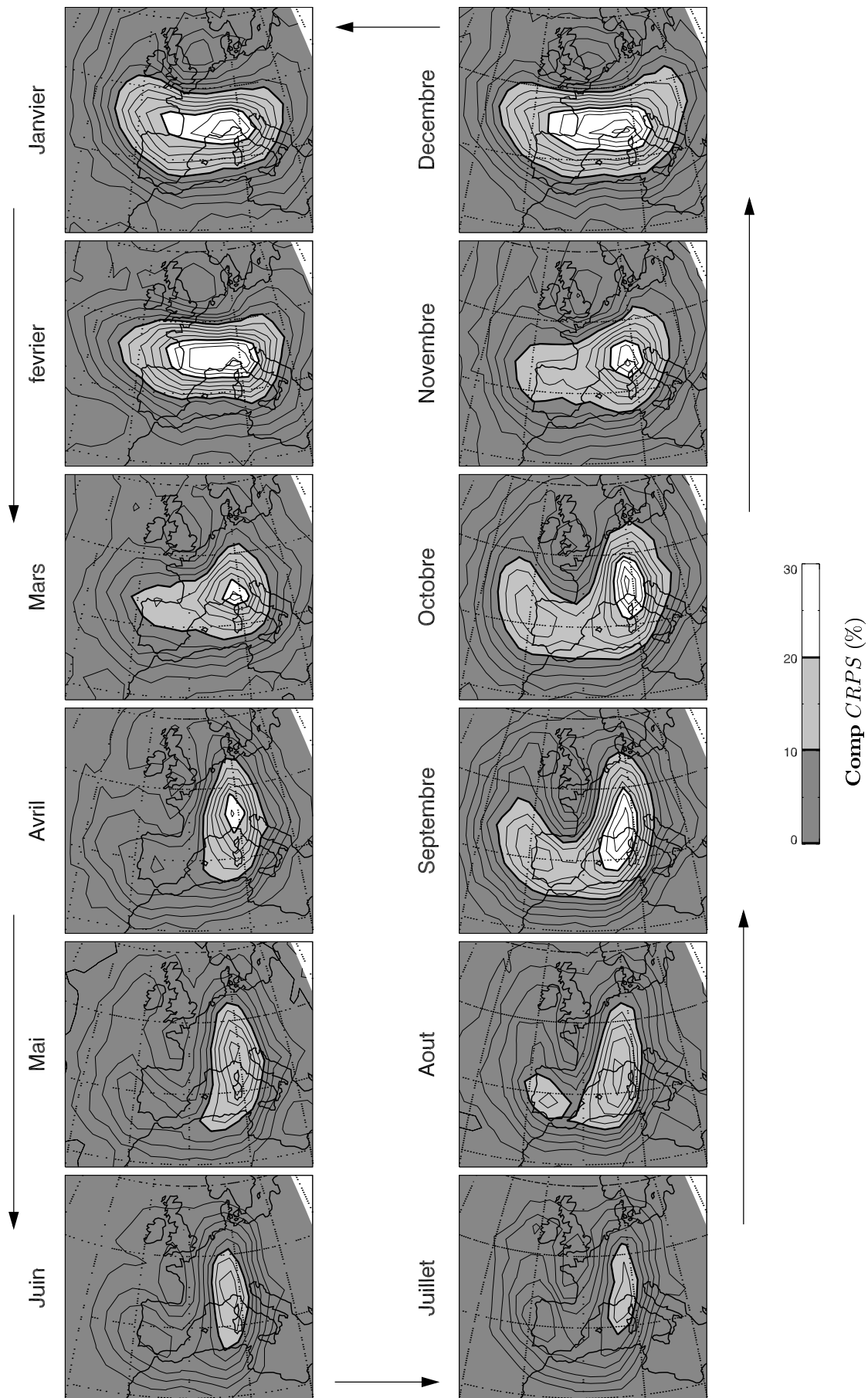


FIG. 5.13: Cartographies de pertinence propres à chaque mois de l'année, groupement Arc_à_Bramans.

alors c'est la fenêtre d'analogie qui aura été optimisée sur le mois d'octobre qui sera utilisée. Néanmoins, même si elle est susceptible de donner des résultats satisfaisants en calibration, cette approche risque fort de s'avérer décevante en validation. Il suffit que l'on rencontre une situation cible atypique de la situation dominante d'un mois donné pour que la fenêtre d'analogie soit complètement inadaptée, étant trop spécifique à une situation qui n'est justement pas la situation cible.

Cette évolution mensuelle est en fait due aux différences, tout au long de l'année, de probabilité d'occurrence des divers types de circulation synoptique. Il nous semble donc que l'approche la plus intéressante serait d'adapter directement la fenêtre sur la situation météorologique cible.

5.2.2 Une fenêtre dynamique ?

Nous allons donc abandonner dans cette sous-section l'usage de fenêtres d'analogie statiques pour adopter des fenêtres dynamiques qui s'adaptent à la circulation atmosphérique propre à chaque situation cible. Elles sont donc à présent différentes d'une situation à une autre.

5.2.2.1 Principe de l'approche dynamique

Lorsque l'on s'intéresse à des cartographies de pertinence pour certaines situations cibles considérées individuellement, telles que celles présentées par la figure 5.14, on s'aperçoit que l'on rencontre les points de grilles pertinents essentiellement en remontant les géopotentiels à partir du groupement d'intérêt. On retrouve donc ici les interprétations précédentes : il semble qu'une fenêtre d'analogie optimale doive considérer principalement les éléments qui arrivent sur le groupement.

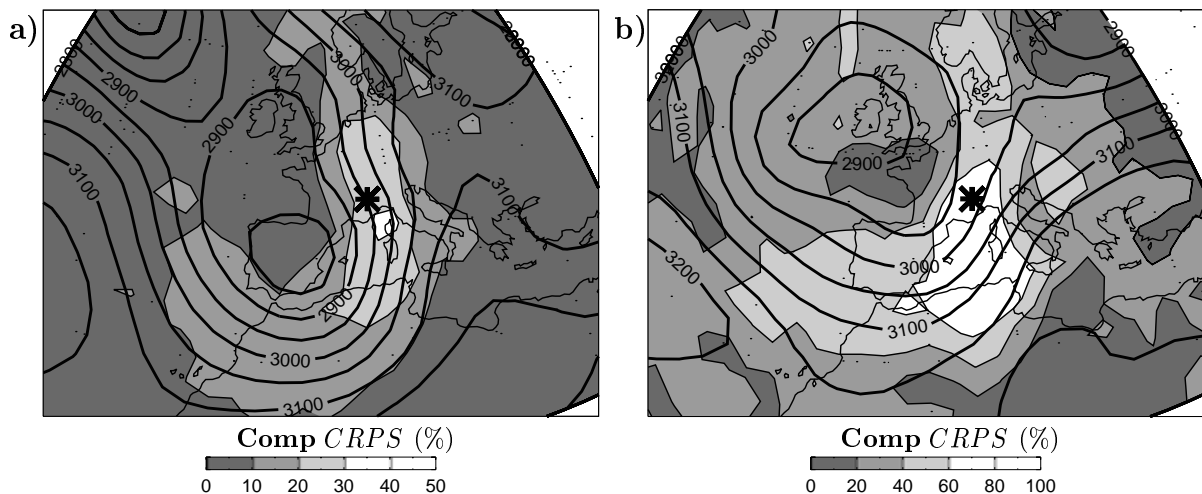


FIG. 5.14: Géopotentiels $z_{700+12h}$ (courbes de niveau) et cartographies de pertinence (contours pleins) sur le groupement *Arc_à_Bramans* (étoile sur la carte) pour deux situations cibles différentes : **a)** le 7 novembre 1962 ; **b)** le 25 octobre 1960.

Nous allons donc construire des fenêtres dynamiques en nous basant sur les géopotentiels de la situation cible. En partant de la fenêtre élémentaire centrée sur le groupement d'intérêt, nous construisons ainsi une fenêtre dynamique selon 4 paramètres :

- **LAMONT** : longueur en remontant le vent géostrophique (*i.e.* en remontant un même géopotentiel),
- **LAVAL** : longueur en descendant le vent géostrophique,
- **THP** : longueur selon une direction perpendiculaire au vent géostrophique, en allant vers les hautes pressions,
- **TBP** : longueur selon une direction perpendiculaire au vent géostrophique, en allant vers les basses pressions.

Tous ces paramètres sont mesurés en *nombre de points de grille*. Les fenêtres ainsi formées ne sont bien sûr pas nécessairement rectangulaires.

En pratique, il faut tout d'abord choisir le géopotentiel qui servira de *base* à la construction de la fenêtre dynamique. Nous raisonnons ensuite en terme de « pavés » de 3x3 points de grille, en procédant de la façon suivante :

- ① Nous sélectionnons tout d'abord le pavé centré sur le groupement qui nous intéresse (figure 5.15a).
- ② Afin de se diriger perpendiculairement au vent géostrophique en direction des basses pressions, nous recherchons quel point de ce pavé présente le plus grand écart négatif de géopotentiel avec le point central. Ce point, que nous appelons MBP, forme alors le centre d'un nouveau pavé. Si $TBP > 1$, cette opération est renouvelée à partir du pavé nouvellement créé autant de fois que nécessaire.
- ③ Nous nous dirigeons ensuite toujours perpendiculairement au vent géostrophique, mais en direction des hautes pressions maintenant. Nous recherchons donc le point MHP du pavé de départ qui présente le plus grand écart positif de géopotentiel avec le point central, afin qu'il forme le centre d'un nouveau pavé (figure 5.15b). De même que précédemment, Si $THP > 1$, cette opération est renouvelée à partir du pavé nouvellement créé autant de fois que nécessaire.
- ④ A la suite de cela, nous remontons les géopotentiels. Parmi les 9 points situés autour du point central du pavé initial, nous recherchons donc celui qui, en allant de MBP à MHP selon le sens trigonométrique, présente le moins d'écart de géopotentiels avec le point central. Ce point définit le centre d'un nouveau pavé, situé en amont par rapport à l'écoulement du vent géostrophique. A partir de ce pavé, nous reconduisons les opérations ② et ③ autant de fois que nécessaire (figure 5.15c). Cette étape est renouvelée $TAMONT$ fois.
- ⑤ Enfin, nous recommençons l'opération ④ à $TALVAL$ reprises, mais en nous plaçant cette fois en aval du vent géostrophique (figure 5.15d).

Même si les fenêtres dynamiques sont différentes d'une situation cible à l'autre, les paramètres qui servent à les construire sont eux toujours les mêmes. Ils se substituent en quelque sorte aux paramètres $lon0$, $lat0$, $dlon$ et $dlat$ des fenêtres statiques.

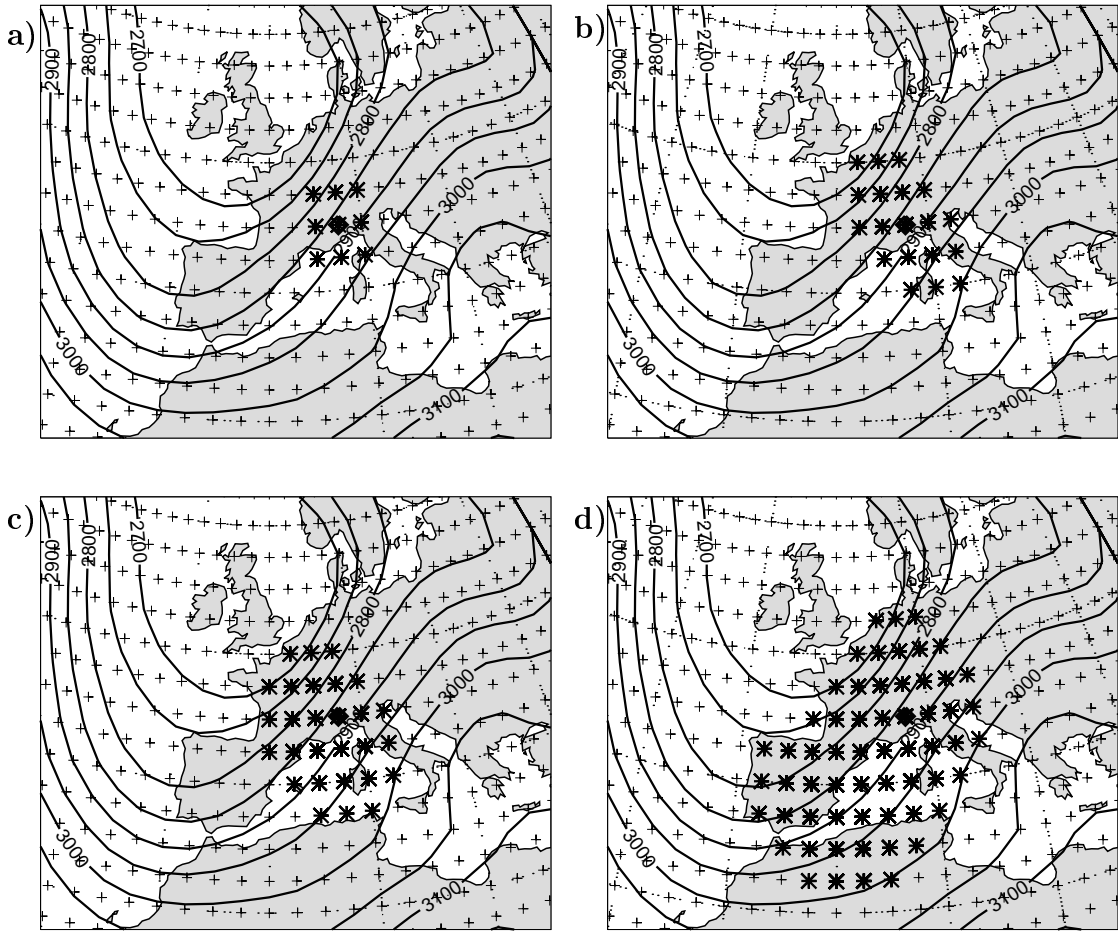


FIG. 5.15: Exemple de construction d'une fenêtre dynamique (étoiles sur les cartes) pour la situation cible du 9 décembre 1954, à partir du géopotentiel $z_{700}+12h$. Le groupement considéré est représenté par un losange sur les cartes. **a)** pavé de départ ; **b)** LAMONT=0, LAVAL=0, THP=1 et TBP=1 ; **c)** LAMONT=1, LAVAL=0, THP=1 et TBP=1 ; **d)** LAMONT=4, LAVAL=1, THP=1 et TBP=1.

Pour définir ces paramètres, nous procédons par calibration, de manière analogue à la façon dont nous avons procédé pour les fenêtres statiques. A partir du pavé initial, nous testons l'agrandissement de chaque fenêtre étape selon l'un des 4 paramètres LAMONT, LAVAL, THP ou TBP. La fenêtre agrandie qui présente les meilleures performances devient fenêtre étape et nous recommençons alors l'opération. Finalement, les paramètres correspondant à la fenêtre étape optimale sont ceux retenus.

5.2.2.2 Résultats obtenus

Il est clair que la définition de la fenêtre dynamique dépend grandement du groupement considéré. Il serait absurde d'en construire une en s'intéressant à un groupement alpin pour ensuite aller l'utiliser avec un groupement pyrénéen. Les fenêtres dynamiques ne se conçoivent donc que comme des fenêtres individuelles ou des fenêtres uniques sur un regroupement, ce qui implique une augmentation significative des temps de calcul. Nous nous abstenons donc de

présenter ici des résultats concernant l'ensemble des groupements, pour nous concentrer sur le seul groupement **Doux-Eyrieux**, choisi arbitrairement.

Choix du point de départ

Pour définir le premier pavé constitutif de la fenêtre dynamique, deux approches différentes sont envisageables :

- La première possibilité est de considérer simplement le point de grille M_{geo} le plus proche du groupement, géographiquement parlant.
- L'autre approche consiste à se dire que ce point M_{geo} , même s'il est le plus proche, n'est pas nécessairement le plus représentatif du groupement, du fait de la climatologie locale ou des conditions topographiques (groupement et point M_{geo} placés de part et d'autre d'une ligne de crête par exemple). On peut alors chercher un point caractéristique M_{carac} à partir de la cartographie de pertinence¹ établie pour ce groupement, en retenant le point qui se montre le plus pertinent.

Des tests effectués sur de nombreux groupements ont montré que les points M_{geo} et M_{carac} étaient effectivement distincts dans la plupart des cas, comme l'illustre la figure 5.16 pour le groupement **Doux-Eyrieux**.

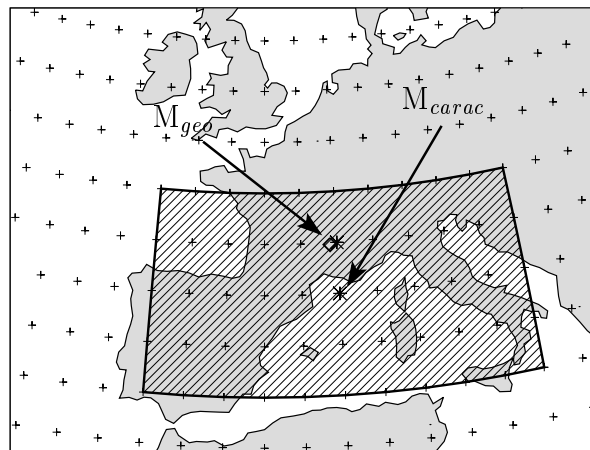


FIG. 5.16: Représentation de la fenêtre unique individuelle optimale (géopotential $z850+12$, rectangle hachuré) et des points M_{geo} et M_{carac} (étoiles) associés au groupement **Doux-Eyrieux** (losange).

En considérant la variable d'analogie $z850+12$, nous avons comparé les résultats obtenus avec d'une part l'utilisation d'une fenêtre unique individuelle optimale, et d'autre part des fenêtres dynamiques optimales construites sur $z850+12$ à partir de M_{geo} ou de M_{carac} (tableau 5.3). Il apparaît que l'utilisation du point caractéristique est plus performante que celle du point le plus proche géographiquement : c'est celui que nous allons donc utiliser dans toute la suite.

On constate par ailleurs que la direction privilégiée par les fenêtres dynamiques est effectivement LAMONT, ce qui confirme que dans le cas général, il convient d'orienter la fenêtre d'analogie vers la provenance du flux qui intéresse le groupement. Néanmoins, les fenêtres dyna-

¹Pour cette application, les cartographies de pertinence sont établies à l'aide de fenêtres élémentaires de 3×3 points, afin que leur barycentre corresponde bien à un point de grille.

FENÊTRES DYNAMIQUES					
Point de départ	LAVAL	LAMONT	THP	TBP	Comp CRPS (%)
M_{geo}	1	4	2	1	34.98
M_{carac}	1	4	1	1	35.44
FENÊTRES STATIQUES INDIVIDUELLES					
lon0 (°E)	lat0 (°N)	dlon (°)	dlat (°)	Comp CRPS (%)	
-7.5	37.5	25	10	36.35	

TAB. 5.3: Fenêtres statiques et dynamiques optimales sur le groupement Doux-Eyrieux. Variable d'analogie : z850+12 ; Base des fenêtres dynamiques : z850+12.

miques s'avèrent ici moins utiles qu'une simple fenêtre statique... Peut-être cela provient-il d'un mauvais choix du champ de géopotentiel servant à construire la fenêtre dynamique : le champ de base.

Influence du champ de base

Même si l'on considère toujours le géopotentiel z850+12 comme variable d'analogie, il est possible d'utiliser un autre champ pour guider la construction de la fenêtre dynamique. Nous appelons ce géopotentiel le *champ de base* de la fenêtre dynamique. Afin d'étudier l'importance de son choix, nous avons calculé les performances des fenêtres dynamiques optimales obtenues en fonction :

- de son niveau, à heure de mesure fixée (+12h) : figure 5.17a,
- et de son heure de mesure, à niveau fixé (850 hPa) : figure 5.17b.

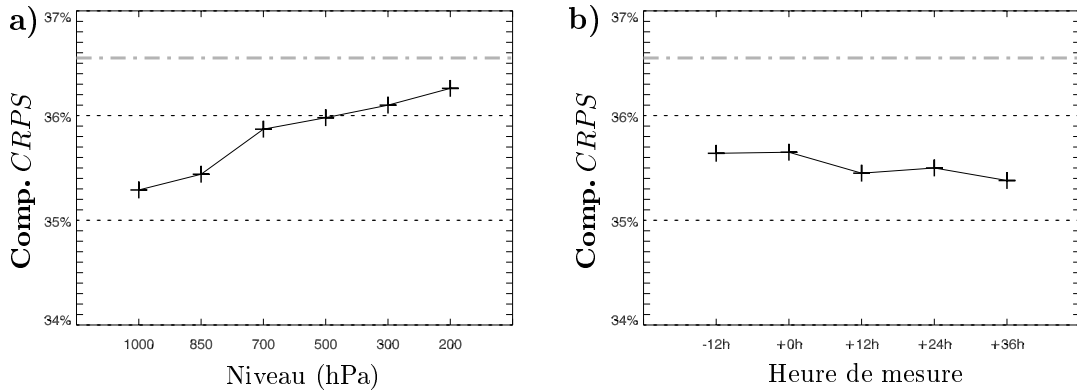


FIG. 5.17: Performances atteintes (**Comp. CRPS**) sur le groupement Doux-Eyrieux par l'utilisation de fenêtre dynamique (variable d'analogie : z850+12) en fonction du champ de base des fenêtres dynamiques. Trait tireté grisé : performance avec une fenêtre statique. **a)** heure de mesure du champ de base fixée à +12h ; **b)** niveau du champ de base fixé à 850 hPa.

Comme les résultats obtenus le montrent clairement, l'utilisation de fenêtres dynamiques se révèle systématiquement moins performante que celle d'une simple fenêtre statique. Le choix du champ de base permet certes d'améliorer les résultats, mais jamais de dépasser ceux obtenus dans un contexte statique, ni même de les égaler.

Il apparaît que plus le niveau du champ de base est élevé, plus les prévisions sont utiles, même si le niveau de la variable d'analogie reste relativement bas (850 hPa). Il y a ainsi un gain de performance notable entre 850 et 700 hPa. Or, c'est effectivement au delà de la couche limite atmosphérique que le vent réel est le plus proche du vent géostrophique. En deçà, les frottements sur la surface terrestre ajoutent une contribution importante sur sa détermination : la fenêtre alors définie s'écarte trop de la circulation réelle pour intégrer les points de grille les plus pertinents. De plus, les niveaux élevés présentent des patterns plus lisses, qui sont donc mieux adaptés à notre algorithme, lequel s'appuie sur une vision assez locale des géopotentiels (seulement 3x3 points de grille à chaque étape).

Etonnamment, l'heure de mesure semble jouer un rôle négligeable. En fait, il s'avère que si l'on s'écarte de +12h, les fenêtres dynamiques optimales sont généralement plus grandes et plus isotropes. Ce que l'on perd donc à cause de l'imprécision sur la circulation effective entre +6 et +30h, on le gagne en considérant un domaine moins spécifique. On se rapproche ainsi d'une fenêtre statique.

Conclusions

Finalement, l'utilisation de fenêtres dynamiques ne s'est guère révélée concluante. Les performances globales obtenues n'arrivent pas à dépasser celles atteintes par une simple fenêtre statique individuelle. Nous n'avons présenté ici que le groupement Doux-Eyrieux, mais cette conclusion est générale pour la dizaine de groupements testée.

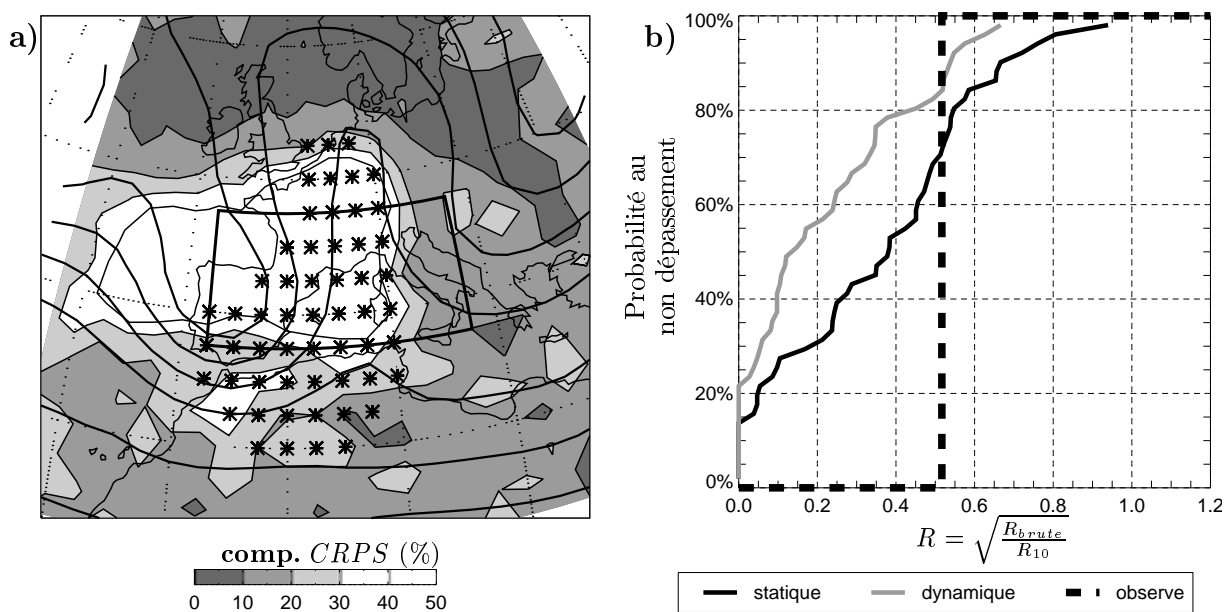


FIG. 5.18: Journée cible du 22/10/1990 pour le groupement Doux-Eyrieux. **a)** : cartographie de la pertinence (contours pleins), géopotentiels z200+12 (courbes de niveau), fenêtre dynamique (étoiles) et fenêtre statique (rectangle) ; **b)** fonction de distribution de la prévision émise avec l'utilisation d'une fenêtre dynamique (courbe noire) et d'une fenêtre statique (courbe grise).

Toutefois, même si elles ne sont pas intéressantes en moyenne, les fenêtres dynamiques permettent d'améliorer substantiellement les prévisions pour certaines situations cibles. Ainsi, la figure 5.18 présente les prévisions émises pour le 22 octobre 1990 par une fenêtre statique et une fenêtre dynamique. Dans ce cas, l'utilisation d'une fenêtre dynamique permet d'éviter de rechercher l'analogie vers l'Est (zone particulièrement peu pertinente pour cette situation) tout en prenant en compte l'information présente au Sud-Est (zone relativement pertinente).

La construction de la fenêtre dynamique est satisfaisante : elle épouse effectivement la forme des géopotentiels. Malheureusement, comme on le voit dans cet exemple, les points de grille pertinents ne suivent pas les géopotentiels de façon aussi exacte. On rencontre ainsi une zone très pertinente dans la dépression située à l'Ouest, que la fenêtre dynamique prend bien soin d'éviter !

Les zones les plus pertinentes ne se situent donc pas systématiquement en amont du vent géostrophique. Les fenêtres dynamiques privilégiant fortement cette direction, leur emploi peut donc s'avérer catastrophique. A l'inverse, même si une fenêtre statique intègre toujours des points de grille non pertinents, elle recouvre une aire suffisamment vaste dans toutes les directions pour que l'on soit sûr de rencontrer dans tous les cas une zone pertinente. Finalement, le gain de performance obtenu par les fenêtres dynamiques lorsque leur spécificité permet de mieux cerner les zones pertinentes reste limité : il n'est pas suffisant pour compenser les cas où elles passent à côté des points de grille réellement pertinents.

5.2.3 Classification par type de temps

L'utilisation de fenêtres dynamiques s'est finalement révélée plutôt décevante. Il est possible que la manière de les construire, en considérant les géopotentiels à des échelles trop locales (pavés de 3x3 points), perde de vue l'échelle synoptique. Nous avons donc testé une autre approche qui se situe à mi-chemin entre l'approche statique et dynamique. Les fenêtres d'analogie utilisées ici dépendent en effet toujours de la circulation atmosphérique de la journée cible, mais sont maintenant basées sur une classification à *échelle synoptique* des situations rencontrées.

5.2.3.1 Méthodologie

Principe des classifications

Les climatologues s'intéressent aux relations statistiques reliant les différentes situations atmosphériques. Cependant, la caractérisation d'une seule situation nécessite la connaissance de la répartition spatiale de toutes les variables atmosphériques. Celles-ci incluent aussi bien les variables dynamiques (pression, vitesse du vent, ...), que les variables thermodynamiques (température, ...) ou encore les variables relatives à l'état hygrométrique (rapport de mélange, ...) : l'ensemble constitue une masse de données considérable. C'est pourquoi les climatologues ont très tôt cherché à agréger les situations atmosphériques en un nombre limité de groupes définis par un nombre restreint de variables. On parle alors de classification en types de temps.

Il existe de nombreuses méthodes de classification (objectives ou subjectives), qui s'appuient généralement soit sur la description de cartes de champs de pression, soit sur la considération des mouvements des masses d'air et des centres d'actions. Ces deux approches ne sont bien sûr pas totalement indissociables, les mouvements des masses d'air étant en grande partie générés par les champs de pression. On peut donc globalement dire qu'un type de temps correspond à un certain type d'organisation spatiale de la circulation générale (Thillet 1997).

Evidemment, il n'existe pas une classification universelle. Chacune dépend de la zone considérée, des variables retenues pour caractériser les situations atmosphériques, de la méthode choisie, du nombre de classes retenues, mais aussi de ce que l'on veut expliquer grâce à cette classification. Le principal inconvénient des classifications est, comme le signalent Barry et Perry (1973), que les modes atmosphériques sont continus, si bien que les définitions de frontières entre les différents groupes sont forcément arbitraires et donc toujours sujettes à insatisfactions. Des techniques avancées de logique floue permettent toutefois de minimiser ce problème, en autorisant une situation à appartenir « *un peu* » à plusieurs types à la fois (Bardossy *et al.* 1995).

Les précipitations ne forment qu'un composant parmi tant d'autres d'une situation atmosphérique donnée. Il est donc fréquent de chercher à les relier aux types de temps fournis par une classification pour construire des générateurs de pluies (voir par exemple Hay *et al.* 1991 ; Stehlik et Bardossy 2002). Plusieurs travaux ont en outre essayé de valoriser ces liaisons dans le cadre de prévisions de précipitations (Clement 1976, appliqué par exemple sur la France).

Approche proposée

L'approche proposée ici consiste à définir tout d'abord un certain nombre de types de temps, au sens de la circulation synoptique, en fonction de leur influence sur les précipitations. Une classification de toutes les situations cibles de l'archive est ensuite réalisée selon ces types de temps : chaque situation cible est rattachée à l'un d'entre eux. Ensuite, toujours dans une phase d'apprentissage, on optimise pour chaque type de temps (en ne considérant que les situations cibles qui lui correspondent), la fenêtre d'analogie qui permet de reconstituer au mieux les pluies observées. Ainsi, à chacun des types de temps définis est finalement associée une fenêtre d'analogie propre.

Dans une application opérationnelle :

- on définira le type de temps associé à la situation cible, et donc la fenêtre d'analogie la plus appropriée,
- mais on donnera leur chance à toutes les situations candidates de l'archive, et non pas seulement à celles qui correspondent au même type de temps que la situation cible.

Le côté « *figé* » de la sélection par type de temps est assoupli par le fait que les analogues sont toujours recherchées selon un critère de plus proche voisin et sans que les situations candidates ne soient restreintes à celles présentant le même type de temps que la situation cible. Seules les fenêtres d'analogies, une pour chaque type de temps, sont figées.

5.2.3.2 Sélection des types de temps

L'objet de ce travail n'est pas de se pencher plus avant sur les techniques de classification en types de temps et encore moins de rechercher une classification optimale pour chacun de nos groupements. Nous nous plaçons ici comme simples utilisateurs de classifications pré-existantes afin d'étudier si leur utilisation peut apporter une amélioration à une adaptation par recherche d'analogies. Si les résultats sont concluants, il sera toujours temps de rechercher des classifications plus pertinentes, en s'appuyant de manière spécifique sur les informations dont nous disposons effectivement.

La classification adoptée

Nous avons donc choisi d'utiliser une classification présentée par Thillet (1997). Etablie de façon subjective, cette classification mêle description des cartes de pression et des trajectoires des masses d'air. Elle vise principalement à expliquer le temps sensible sur la France, et plus particulièrement sur les Alpes. La principale motivation pour son utilisation ici réside dans sa simplicité. Elle ne considère en effet que 6 types de temps distincts pour lesquels les auteurs citent à chaque fois un exemple de situation réelle caractéristique, choisi à l'intérieur des années 1992-94 (figure 5.19) :

- Ⓣ₁ **Régime d'Ouest** : il se caractérise par une zone de hautes pressions centrée sur les Açores qui se prolonge par une dorsale vers l'Europe de l'Ouest, accompagnée par de relativement basses pressions sur l'Atlantique Nord. Situation caractéristique : 16 juin 1993.
- Ⓣ₂ **Régime de Nord-Ouest** : l'anticyclone des Açores s'étend ici vers l'Islande tandis qu'une dépression se situe sur le sud de la Scandinavie, ce qui entraîne un flux de Nord-Ouest souvent perturbé sur la France. Situation caractéristique : 13 mars 1994.
- Ⓣ₃ **Régime de Sud-Ouest** : ce régime présente une importante zone dépressionnaire ou un thalweg sur les îles britanniques, alors qu'une dorsale s'étend généralement du bassin méditerranéen à la Russie. Les perturbations abordent alors la France par le Sud-Ouest. Situation caractéristique : 5 octobre 1993.
- Ⓣ₄ **Retour d'Est** : ce régime naît du creusement d'une dépression sur la méditerranée occidentale, le plus souvent créé par la scission d'un thalweg ayant amené de l'air polaire. Les masses d'air qui circulent autour de cette goutte froide sont généralement douces et instables. Elles se chargent fortement d'humidité au dessus de la Méditerranée avant d'aborder la France par l'Est en étant susceptibles de provoquer des précipitations très intenses sur les reliefs proches des rivages méditerranéens. Situation caractéristique : 5 octobre 1992.
- Ⓣ₅ **Régime d'Est** : il se caractérise par une zone de hautes pressions au dessus de la Russie qui amène des masses d'air continentales sèches et froides au dessus de la France. Cette situation est souvent à l'origine des vagues de froids. Situation caractéristique : 1er janvier 1993.
- Ⓣ₆ **Régime anticyclonique** : comme son nom l'indique, lors de cette situation, un anticyclone est ancré au dessus de l'Europe de l'Ouest. Généralement signe de beau temps, cette situation n'empêche pas la formation d'orages isolés, notamment sur les reliefs. Situation caractéristique : 3 février 1993.

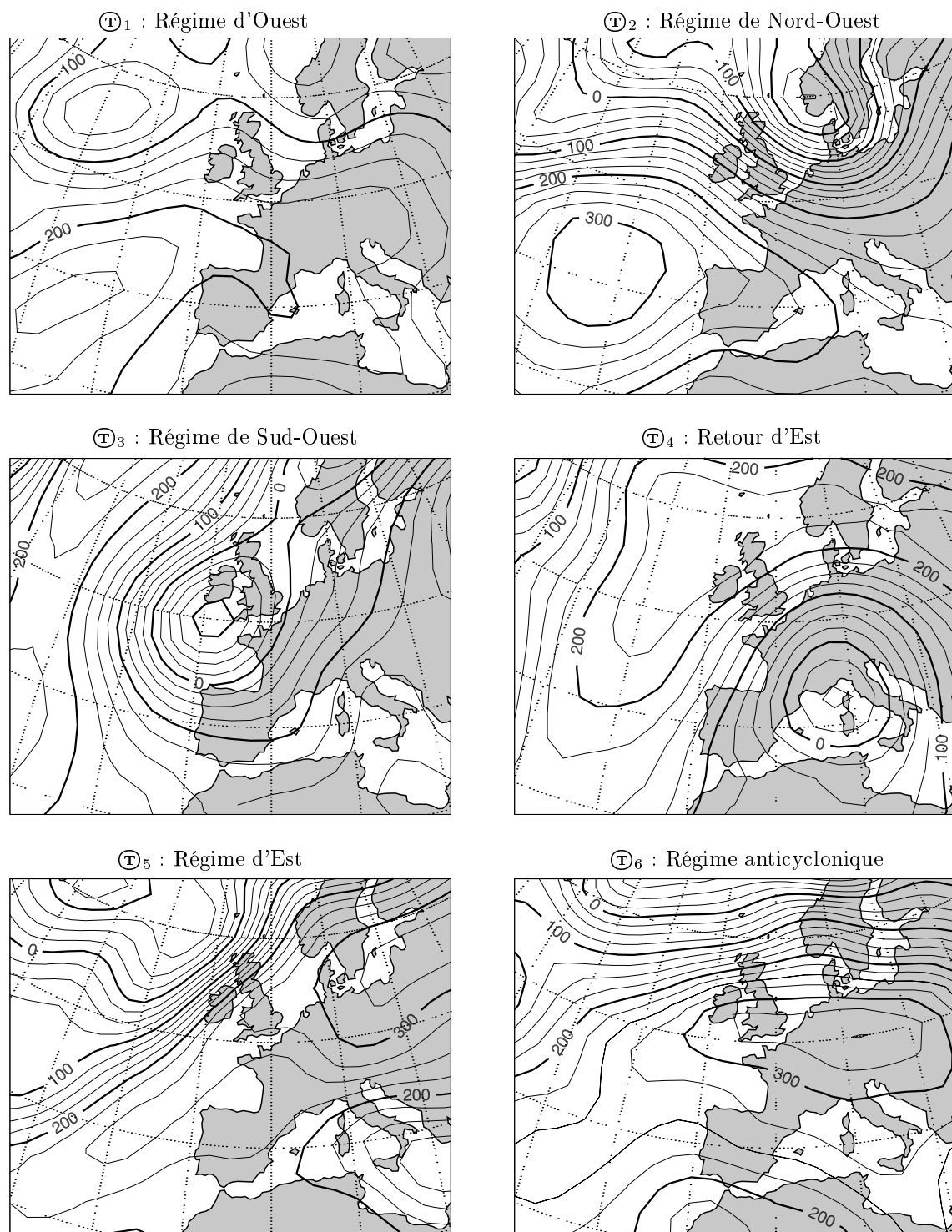


FIG. 5.19: Géopotentiels $z_{1000}+12h$ des situations caractéristiques de chacun des types de temps retenus.

Pour associer l'un de ces types de temps à chacune des journées cibles, nous ne sommes pas en mesure de considérer de façon systématique les trajectoires des masses d'air associées à chacune d'entre elles. Nous avons donc choisi de nous contenter de l'information contenue dans les seuls champs de pression. Ce choix nous permet de procéder simplement par analogie, en considérant pour chaque situation cible, les situations candidates formées par les 6 situations caractéristiques. La mesure de l'analogie entre la situation cible et chacune de ces 6 situations permet de les classer de la plus similaire à la moins similaire. La journée courante sera finalement associée au type de temps correspondant à la situation caractéristique qui présente la plus forte analogie.

Cette analogie porte sur les champs de géopotentiels ($z_{1000+12}$), comparés par l'intermédiaire du score $S1$, étant donné que c'est la circulation atmosphérique que nous cherchons à caractériser via la classification en type de temps. La fenêtre d'analogie adoptée correspond à la fenêtre statique optimisée pour l'utilisation d'un seul champ de géopotentiel (z_{850+12}) : $d_{lon}=27.5^\circ$ et $d_{lat}=10^\circ$.

En procédant ainsi, nos situations cibles successives sont réparties en 6 groupes (figure 5.20). Comme nous l'évoquions précédemment (cf. 5.2.1), on constate bien qu'elles se répartissent de façon inégale selon les mois de l'année. Ainsi, les Régimes de Sud-Ouest, qui sont les situations les plus fréquentes (22%), se rencontrent de façon à peu près constante tout au long de l'année, sauf durant les mois de juin-juillet-août où elles se font deux fois plus rares. Il est intéressant de noter par ailleurs que les Régimes d'Ouest, assez peu présents en hiver, dominent très largement les circulations des mois de juillet et d'août, et que ce sont les Retour d'Est qui prédominent

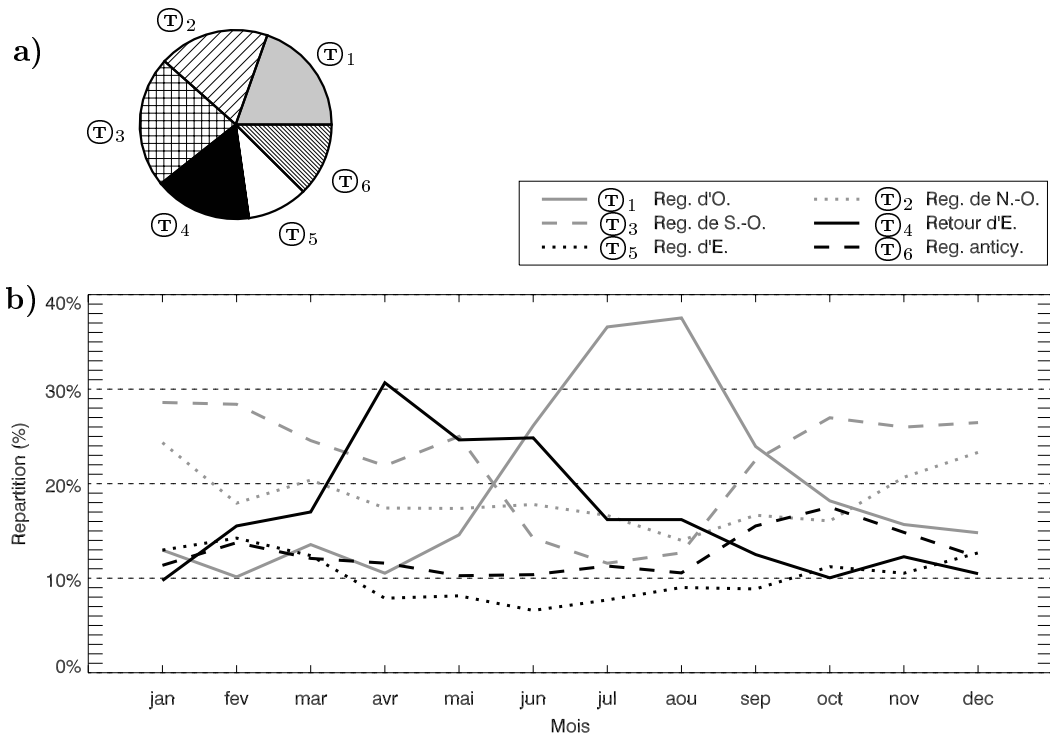


FIG. 5.20: Répartition des différents types de temps au sein des situations cibles, **a)** : répartition inter-annuelle; **b)** : répartition mensuelle.

au printemps (avril-mai). Néanmoins, aucun des types de temps n'est totalement absent d'une saison donnée, ce qui justifie que nous n'ayons pas procédé à une simple saisonnalisation des algorithmes.

Pouvoir discriminant de cette classification

Il apparaît que cette classification, bien que non optimisée, possède déjà un réel pouvoir discriminant au niveau des précipitations. La figure 5.21 présente ainsi les distributions empiriques de précipitations associées à chaque type de temps sur le groupement **Fier**. A chaque fois, elles s'écartent notablement de la distribution climatologique (*i.e.* toute situations confondues). Ainsi, les régimes d'Est et anticycloniques présentent de faibles probabilités d'occurrence de pluie (inférieures à 20%), alors que le régime de Nord-Ouest est particulièrement propice aux précipitations (probabilité d'occurrence supérieure à 80%). Si les différents types de temps retenus étaient sans influence sur les précipitations, ces distributions ne seraient que des échantillonnages aléatoires de la climatologie et seraient donc confondues avec elle.

Une méthode basique de prévision probabiliste des précipitations pourrait consister simplement, après avoir déterminé le type de temps correspondant à la situation cible, à annoncer la distribution empirique associée à celui-ci. On disposerait ainsi de 6 prévisions différentes possibles. Le calcul de l'utilité de cette prévision, par l'intermédiaire de la compétence du *CRPS* par rapport à la climatologie globale par exemple, permet de quantifier le pouvoir discriminant de la classification par rapport aux précipitations.

Pour le groupement **Fier**, l'utilité **Comp. CRPS** de la classification atteint ainsi **17.41%**. Ce score n'est pas négligeable, mais reste toutefois inférieur à une approche par analogie : l'utilisation d'une fenêtre statique individuelle et des géopotentiels $z_{1000}+12$ et $z_{500}+24$ réalise pour sa part un score de **45.21%**.

Voilà un résultat rassurant : l'adaptation par analogie est effectivement plus efficace qu'une application triviale d'une classification en type de temps !

D'un groupement à l'autre, le pouvoir discriminant de la classification n'est pas identique. On observe de fortes disparités, entre par exemple le groupement **Tech**, pour lequel la classification atteint péniblement un score de 3.50% (contre 28.04 % par analogie) et le groupement **Cure**, qui permet de grimper jusqu'à 19.35% (contre 40.87 % par analogie).

La figure 5.22a, qui présente le score atteint par la climatologie pour chaque groupement, est à comparer avec la figure 5.9c qui présentait les performances obtenues par analogie en considérant des fenêtres individuelles statiques. On constate alors que ces performances sont fortement liées entre elles (figure 5.22b) : on retrouve bien ici que les groupements réagissent plus ou moins directement à la circulation synoptique, soit du fait de particularités climatiques locales, soit à cause de leur constitution hétérogène. Seul les groupements situés le plus au Nord-Ouest de la France (**Cure**, **Saône** et **Bretagne**) sortent du rang : ce sont ceux pour lesquels cette classification est la plus pertinente en regard de leur sensibilité à la circulation synoptique. Cette sensibilité accrue est peut-être due au caractère « ouvert » de ces groupements, qui présentent peu d'effets

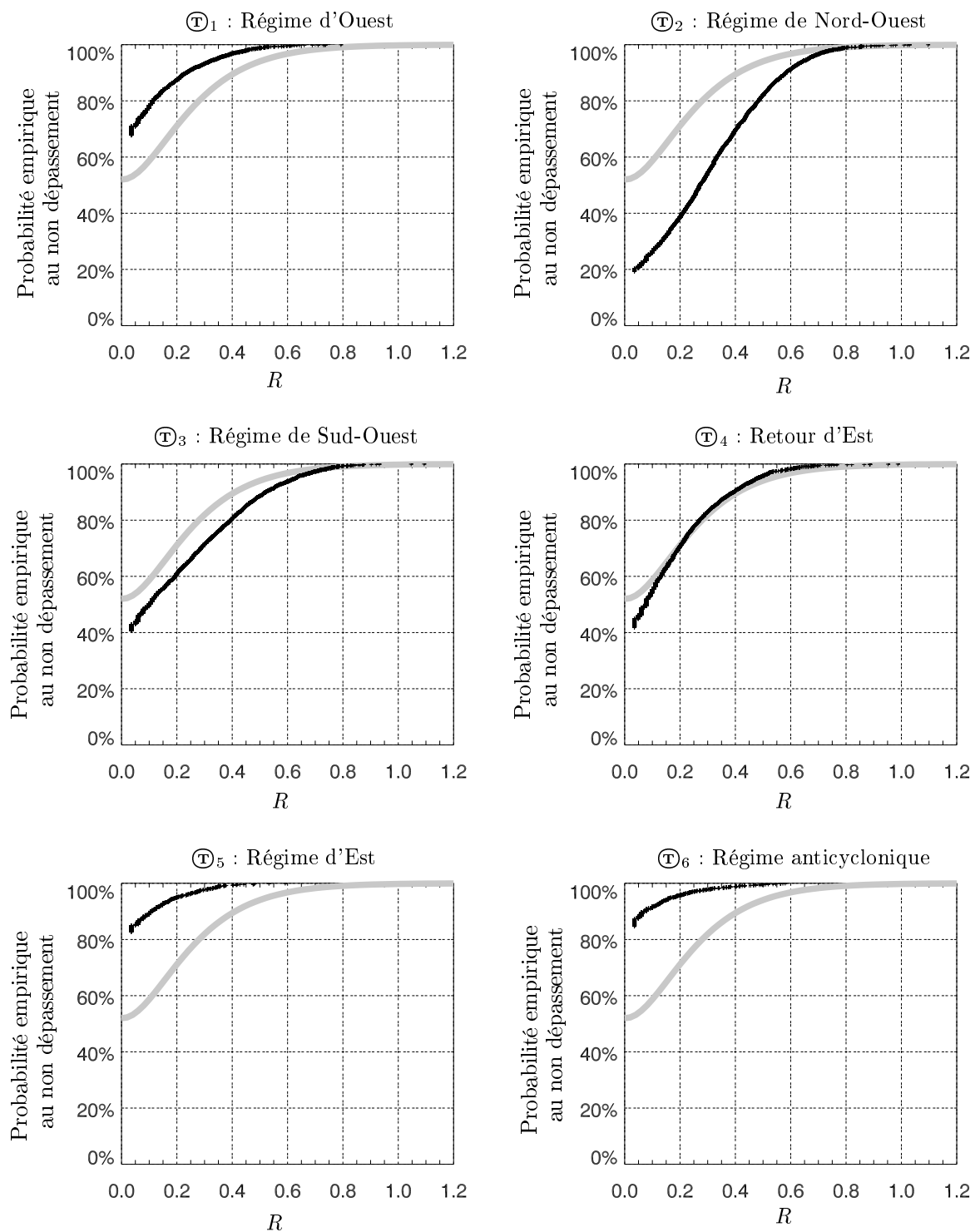


FIG. 5.21: Fonctions de répartition empiriques des précipitations selon chaque type de temps (croix) et de l'ensemble des précipitations (distribution climatologique, courbe grisée) pour le groupement Fier.

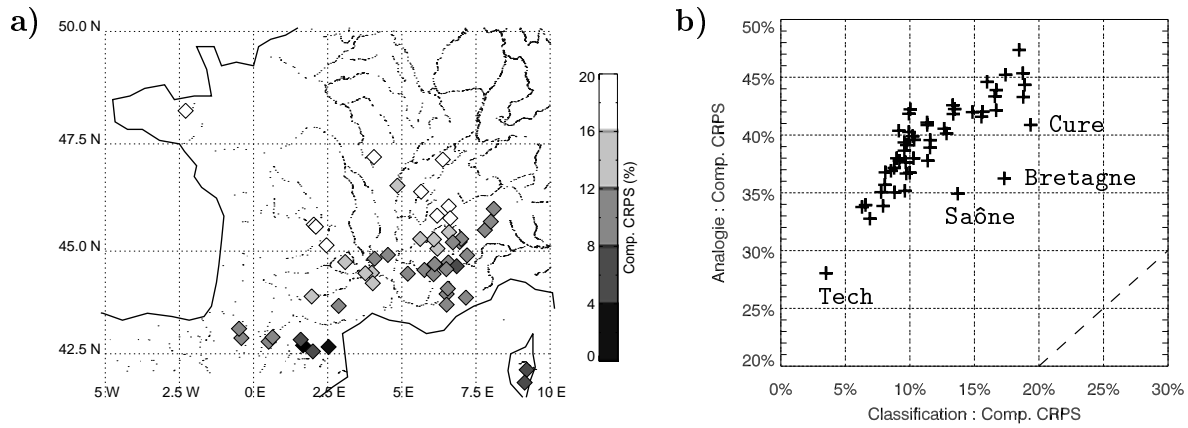


FIG. 5.22: Pouvoir discriminant de la classification en type de temps adoptée vis à vis des précipitations (**Comp. CRPS**) : **a)** cartographie des groupements; **b)** Comparaison des scores en fonction des performances obtenues par analogie.

orographiques susceptibles de dissocier les flux de basse couche de l'organisation synoptique des éléments atmosphériques.

5.2.3.3 Résultats

De même que pour l'approche totalement dynamique considérée précédemment, il n'y aurait pas grand intérêt à utiliser ici des fenêtres communes à l'ensemble des groupements. En effet, chacun d'eux réagit différemment aux divers types de temps. Il n'est donc pas sûr que pour un type de temps donné, tous les groupements doivent s'intéresser à la même fenêtre d'analogie. C'est pourquoi nous présentons les résultats obtenus sur un seul groupement, à savoir le groupement **Fier**, choisi encore une fois de façon plus ou moins arbitraire.

Cartographies de pertinence

Avant de tester l'application d'une sélection de la fenêtre d'analogie basée sur la classification, il est intéressant de considérer les cartographies de pertinence propre à chaque type de temps, présentées par la figure 5.23 pour **Fier**. On remarque alors deux choses importantes :

- D'une part, les niveaux moyens de pertinence sont très variables d'un type de temps à l'autre. Il apparaît par exemple que les précipitations associées aux épisodes de retour d'Est sont très faiblement expliquées par analogie (pertinence maximale inférieure à 20%), alors que celles liées aux régimes anticycloniques ou aux régime d'Est, affichent des pertinences maximales qui dépassent les 50%. Cette disparité ne semble pas liée au caractère plus ou moins sec de tel ou tel type de temps, car les régimes de Nord-Ouest, fortement propices à l'occurrence de précipitations, montrent des pertinences maximales parfaitement honorables, de plus de 40%.
- Les patterns obtenus pour ce groupement ne sont pas tellement différents les uns des autres. Ils prennent tous la forme d'un U plus ou moins couché et plus ou moins centré sur le

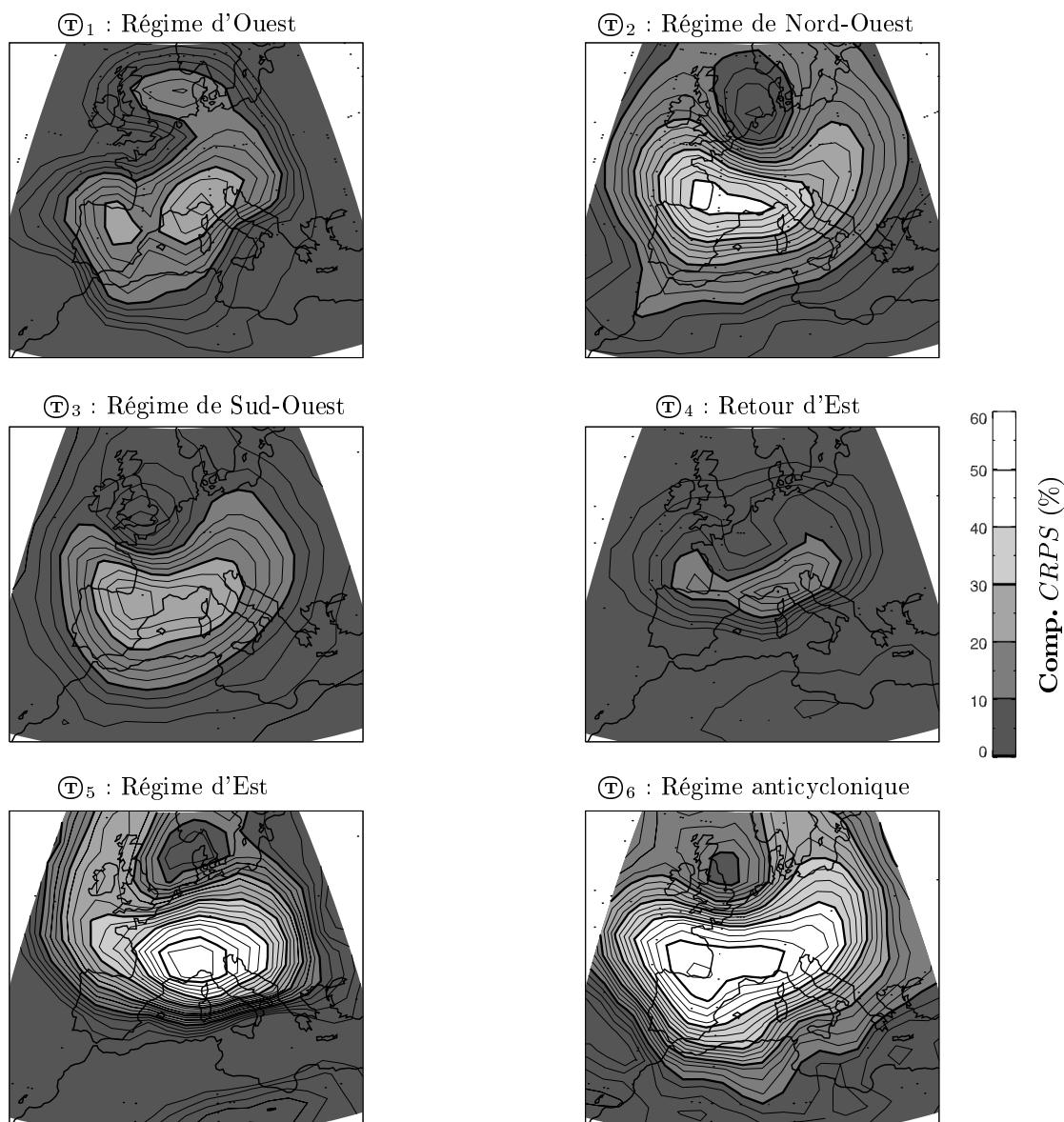


FIG. 5.23: Cartographie de pertinence en fonction du type de temps pour le groupement **Fier**.

groupement. Il s'enroule autour d'une zone de faible pertinence qui se déplace entre la Bretagne et le Danemark, en fonction du type considéré.

Donc, même si, comme on l'a vu, la classification réalise une discrimination certaine sur les précipitations, elle ne semble pas avoir une influence déterminante sur l'organisation spatiale de la pertinence d'analogie. Voilà qui est de mauvais augure pour l'application de fenêtres propres à chaque type de temps...

Application de la classification à la définition de la fenêtre

Le tableau 5.4 présente le résultat de l'application de différents choix de fenêtres d'analogie sur la variable d'analogie ($z_{1000+12}$ & z_{500+24}). Nous considérons ainsi, en plus de la fenêtre

statique indépendante de la classification, 3 ensembles de fenêtres conditionnées par le type de temps de la situation cible :

- des fenêtres statiques mais propres à chaque type : on dispose donc en tout de 6 fenêtres différentes,
- des fenêtres dynamiques dont les paramètres dépendent du type, mais qui considèrent toutes le même point de départ : le point caractéristique du groupement **Fier** sur l'ensemble des situations (champ de base : z200+12),
- et enfin des fenêtres dynamiques dont les paramètres *et* le point de départ sont fonction du type de temps : on utilise alors le point caractéristique associé à chaque type (champ de base : z200+12).

FENÊTRE STATIQUE INDÉPENDANTE						
Type	lon0 (°E)	lat0 (°N)	dlon (°)	dlat (°)	Comp CRPS (%)	
\mathbb{T}_1	-5	40	20	7.5	36.78	
\mathbb{T}_2	-5	40	20	7.5	49.89	
\mathbb{T}_3	-5	40	20	7.5	40.04	
\mathbb{T}_4	-5	40	20	7.5	29.15	
\mathbb{T}_5	-5	40	20	7.5	60.27	
\mathbb{T}_6	-5	40	20	7.5	62.40	

FENÊTRES STATIQUES PAR TYPE						
Type	lon0 (°E)	lat0 (°N)	dlon (°)	dlat (°)	Comp CRPS (%)	Gain (%)
\mathbb{T}_1	-5	40	20	12.5	37.08	+ 0.82
\mathbb{T}_2	-2.5	40	15	7.5	50.07	+ 0.34
\mathbb{T}_3	-5	40	20	7.5	40.04	0.00
\mathbb{T}_4	-7.5	40	25	10	30.12	+ 3.33
\mathbb{T}_5	0	40	12.5	12.5	60.94	+ 1.11
\mathbb{T}_6	-5	37.5	25	10	63.20	+ 1.28

FENÊTRES DYNAMIQUES : POINT DE DÉPART COMMUN						
Type	LAVAL	LAMONT	THP	TBP	Comp CRPS (%)	Gain (%)
\mathbb{T}_1	1	3	1	2	35.25	- 4.16
\mathbb{T}_2	2	3	1	0	49.15	- 1.50
\mathbb{T}_3	1	3	1	1	38.83	- 4.15
\mathbb{T}_4	1	2	1	1	28.73	- 1.44
\mathbb{T}_5	1	1	0	1	58.79	- 2.46
\mathbb{T}_6	2	2	1	2	60.68	- 2.75

FENÊTRES DYNAMIQUES : POINT DE DÉPART FONCTION DU TYPE DE TEMPS						
Type	LAVAL	LAMONT	THP	TBP	Comp CRPS (%)	Gain (%)
\mathbb{T}_1	1	3	1	2	35.77	- 2.75
\mathbb{T}_2	3	1	1	1	48.02	- 3.75
\mathbb{T}_3	2	0	1	2	37.74	- 5.74
\mathbb{T}_4	1	3	1	0	29.32	+ 0.58
\mathbb{T}_5	1	1	0	1	58.79	- 2.46
\mathbb{T}_6	3	0	1	2	59.69	- 4.34

TAB. 5.4: Paramètres et performances des fenêtres optimales associées à chaque type de temps. Le gain correspond à la différence relative de **Comp. CRPS** par rapport aux performances obtenues sur chaque type de temps par l'utilisation d'une fenêtre statique indépendante de la classification (groupement **Fier**).

La différenciation des fenêtres statiques en fonction du type de temps apporte un gain qui dépend du type considéré : il est ainsi nul pour \mathbb{T}_3 (Régime de Sud-Ouest) mais atteint 3.33 % pour \mathbb{T}_4 (Retour d'Est). Cependant, il reste globalement très faible : la taille des fenêtres utilisées fait qu'une légère modification de celles-ci n'a qu'une influence limitée sur les prévisions. La fenêtre statique indépendante constitue donc un compromis qui, pour l'ensemble des situations, ne s'avère pas si mauvais que cela.

Les résultats obtenus avec l'utilisation de fenêtres dynamiques sont plus que médiocres : ils sont mauvais. Quel que soit le type de temps considéré, ils restent inférieurs à ceux d'une fenêtre statique indépendante. Les fenêtres dynamiques engendrent des pertes de performances qui peuvent même atteindre plus de 4 % (\mathbb{T}_1 et \mathbb{T}_3).

La différenciation du point de départ en fonction du type de temps permet parfois de réduire la perte de performance. Cependant, même lorsqu'elle permet d'obtenir un gain de performance (+ 0.58 % pour \mathbb{T}_4), celui-ci reste inférieur à celui obtenu par l'utilisation de simples fenêtres statiques par types de temps. En outre, pour certains types de temps, les points caractéristiques sont très éloignés du groupement. Ils constituent certes une zone très informative dans le cas général, mais la construction de fenêtres dynamiques à partir de ces points ne mène pas systématiquement à considérer un flux qui intéresse effectivement le groupement cible. Les résultats se dégradent alors encore plus, comme c'est le cas avec les types \mathbb{T}_2 , \mathbb{T}_3 et \mathbb{T}_6 .

En conclusion, alors que la classification apporte une réelle discrimination sur les précipitations, la différenciation de la fenêtre d'analogie en fonction du type de temps n'est pas un facteur d'amélioration de l'analogie sur la circulation synoptique. Les différences de cartographie de pertinence d'un type à l'autre sont peu prononcées, si bien que l'utilisation d'une fenêtre statique se révèle finalement suffisamment polyvalente. A partir du moment où celle-ci n'est pas trop petite, la prise en compte de zones peu pertinentes affecte assez peu les performances. Nous retrouvons ici les conclusions précédentes (cf. 5.1.3.1) : il est préférable de considérer une fenêtre plutôt trop grande que trop petite.

La classification n'est toutefois pas dénuée d'intérêt. Elle permet en effet d'estimer a priori, pour chaque situation cible, les performances de l'adaptation par analogie. Ainsi, une fois que l'on a reconnu à quel type de temps appartient une situation cible, on peut se référer aux performances obtenues en calibration pour ce type particulier. De cette manière, on sait si l'on se trouve dans une situation pour laquelle l'adaptation par analogie est en général efficace, ou s'il demeure une forte incertitude quant aux prévisions obtenues.

5.3 Conclusion

Nous avons vu dans ce chapitre que la réalisation de cartographies de pertinence permet de mieux comprendre l'organisation spatiale de l'analogie. Il est ainsi clairement apparu que les choix de la fenêtre d'analogie et des géopotentiels utilisés ne sont pas indépendants : l'heure d'observation influe fortement sur la localisation spatiale de l'information pertinente. A posteriori,

pour optimiser ces choix, la meilleure démarche semble être, comme nous l'avons plus ou moins fait, de :

- ① Choisir a priori une fenêtre d'analogie de taille moyenne ($\approx 20 \times 12^\circ$) centrée sur les groupements cibles.
- ② Rechercher sur cette fenêtre le meilleur couple de géopotentiels. Si ceux-ci ne se situent pas à l'intérieur de la période de cumul des précipitations, reprendre l'étape ① pour être mieux centré sur les groupements.
- ③ A partir de ce couple de géopotentiels, effectuer une optimisation de la fenêtre d'analogie en partant de la fenêtre élémentaire la plus à l'aplomb du centre de gravité des groupements.

L'utilisation d'une seule fenêtre ne peut évidemment pas être optimale pour tout un ensemble de groupements dispersés. De même, il peut être sur-optimal de considérer une fenêtre différente pour chacun des groupements pluviométriques cibles. Nous avons donc établi quelques règles, basées uniquement sur la localisation géographique des groupements, qui permettent de trouver un compromis et d'établir les regroupements adéquats en fonction de l'incertitude tolérée.

Nous avons par ailleurs essayé d'effectuer un conditionnement des fenêtres d'analogies en fonction de la circulation synoptique propre à chaque situation cible. Il est en effet apparu que dans le cas général, la zone la plus pertinente se trouvait en amont du flux qui aborde le groupement cible. Les résultats obtenus ont toutefois été décevants.

Ainsi, le premier conditionnement testé, basé sur une vision locale de la circulation et la construction de fenêtres différentes pour chacune des situations cibles, n'a même pas réussi à égaler les résultats obtenus avec une simple fenêtre statique. En effet, même si la zone pertinente se situe généralement en amont du flux, ce n'est pas toujours le cas : les fenêtres alors utilisées sont totalement inadaptées, de même que les prévisions obtenues. Les pertes de performances associées annulent finalement les gains que le conditionnement peut apporter dans les situations plus standard.

Le second conditionnement testé, qui repose sur une classification en type de temps et l'association à chacun d'eux une fenêtre statique spécifique, apporte pour sa part de très légers gains. Les différences entre les diverses fenêtres spécifiques ne sont effectivement pas suffisamment importantes pour que celles-ci s'avèrent incompatibles avec une fenêtre statique moyenne : l'utilisation d'une fenêtre statique indépendante du type de temps constitue un compromis tout à fait acceptable. Le principal intérêt de la classification réside dans sa capacité à indiquer l'efficacité moyenne de l'analogie vis à vis de chaque type de situation avant même que les événements n'aient lieu. Il est ainsi possible d'associer a priori une incertitude sur la qualité future de la prévision émise.

En conclusion, il semble bien que l'on rencontre une limite sur les performances que l'on peut espérer obtenir grâce à une analogie basée sur la seule circulation synoptique. Nous l'évoquions déjà précédemment, en constatant les faibles gains obtenus en combinant ensemble plus de deux géopotentiels. Le palier atteint par les fenêtres pour ces différentes combinaisons ainsi que l'inefficacité du conditionnement des fenêtres par la circulation propre à la situation cible, confirment à nouveau cette impression.

La partie encore inexpliquée des précipitations ne peut donc a priori plus être recherchée dans la circulation synoptique, quels que soient les raffinements apportés. Pour aller plus loin, il est sûrement nécessaire de prendre en compte de nouvelles variables à caractère plus local, afin d'enrichir l'analogie.

6

Variables de second ordre

**« Tel brille au second rang
qui s'éclipse au premier. »**

Voltaire, La Henriade.

Nous avons vu précédemment que l'analogie sur la circulation synoptique, basée sur la comparaison des géopotentiels, permet déjà d'obtenir des performances intéressantes. Elle semble néanmoins atteindre des limites que l'on n'arrive pas à dépasser, quel que soit le degré de raffinement que l'on mette en œuvre. La partie des précipitations expliquée par la circulation synoptique est donc loin d'être négligeable, mais elle n'est pas forcément suffisante.

Si l'on veut augmenter l'utilité de l'adaptation par analogie, il faut donc inclure dans l'analogie des informations supplémentaires, affinant la caractérisation des situations météorologiques vis à vis des précipitations. Nous devons ainsi ne pas nous contenter des seules données dynamiques, mais pourquoi pas, prendre aussi en compte des données thermiques ou hygrométriques. L'objet de ce chapitre est donc d'étudier ce que peut apporter l'adjonction de variables autres que la circulation synoptique dans une recherche d'analogie. Nous parlerons alors de *variables de second ordre*.

Dans un premier temps, nous allons définir une stratégie qui nous permettra de tenir compte de deux variables de nature différente, afin d'être en mesure de reconnaître des situations analogues à la fois au niveau de la circulation générale ainsi qu'à celui de la variable de second ordre choisie. Cette étape sera l'occasion d'effectuer un premier tri parmi l'ensemble des variables de second ordre potentielles contenues dans l'archive LTHE, toutes n'étant bien sûr pas aussi informatives les unes que les autres vis à vis des précipitations.

Nous retiendrons ensuite la variable de second ordre la plus pertinente sur l'ensemble des groupements pour définir les paramètres d'analogie qu'il convient de lui associer. Enfin, nous testerons une approche par type de temps pour guider le choix de la variable de second ordre.

6.1 Stratégie adoptée

6.1.1 Définition d'une analogie en deux temps

Le principal problème qui se pose à nous pour intégrer de nouvelles données dans l'algorithme d'analogie, est que l'on doit alors comparer entre elles des variables de nature différente. Peut-on par exemple évaluer si, au niveau de l'effet sur les précipitations, 10 mètres de différence sur une hauteur géopotentielle sont équivalents à 3% d'écart sur des humidités relatives ?

En outre, les différentes variables d'analogie possibles ne requièrent pas toutes le même critère d'analogie. Nous avons ainsi vu dans le chapitre 4 que l'utilisation du score $S1$ n'est pertinente qu'avec les géopotentiels, les autres variables préférant généralement un critère de distance moyenne, tel que le $RMSE$.

Une première approche, qui consisterait à rechercher *dans le même temps*, et surtout dans un même critère, une analogie sur deux variables différentes, ne paraît donc pas chose aisée. Il faudrait être capable soit de définir un critère sans dimension commun, soit de pondérer les critères propres à chaque variable. Il est certes possible d'essayer d'obtenir des coefficients de pondération adéquats de manière heuristique, mais cela risque d'être extrêmement gourmand en temps de calcul. En outre, pour chaque nouvelle variable essayée, il sera nécessaire de rechercher de nouveaux coefficients.

C'est pourquoi nous préférons adopter ici une stratégie d'analogie basée sur une sélection en deux temps des situations analogues, comme schématisé par la figure 6.1 :

- ① Dans un premier temps, à partir de l'ensemble des situations de l'archive météorologique, on procède à une sélection d'analogues au sens de la circulation synoptique. On se retrouve ainsi dans le cadre des chapitres précédents, lors desquels nous considérons les géopotentiels comme variables d'analogie et le score $S1$ comme critère d'analogie. Au terme de cette première étape, $N1$ analogues sont identifiées.
- ② Dans un second temps, on ne considère plus que les $N1$ situations retenues lors de la première étape. On les reclasse alors entre elles en recherchant une analogie vis à vis de la situation cible, maintenant basée sur une autre variable d'analogie (la *variable de second ordre*) et un critère d'analogie adapté, différent de $S1$ si besoin est. On ne retient finalement que les $N2$ situations présentant la plus forte analogie au sens de la variable de second ordre.

Cette approche introduit explicitement une hiérarchie dans les analogies. On considère ainsi que l'analogie sur la circulation générale est la plus importante : l'analogie sur la variable de second ordre est significative uniquement si cette première analogie est présente.

Cette façon hiérarchique de procéder nous permet d'utiliser lors de la seconde sélection des fenêtres et des critères d'analogie spécifiques à la variable de second ordre. Il est ainsi possible de considérer une variable qui n'a qu'une influence très locale, comme un taux d'humidité ou un indice d'instabilité.

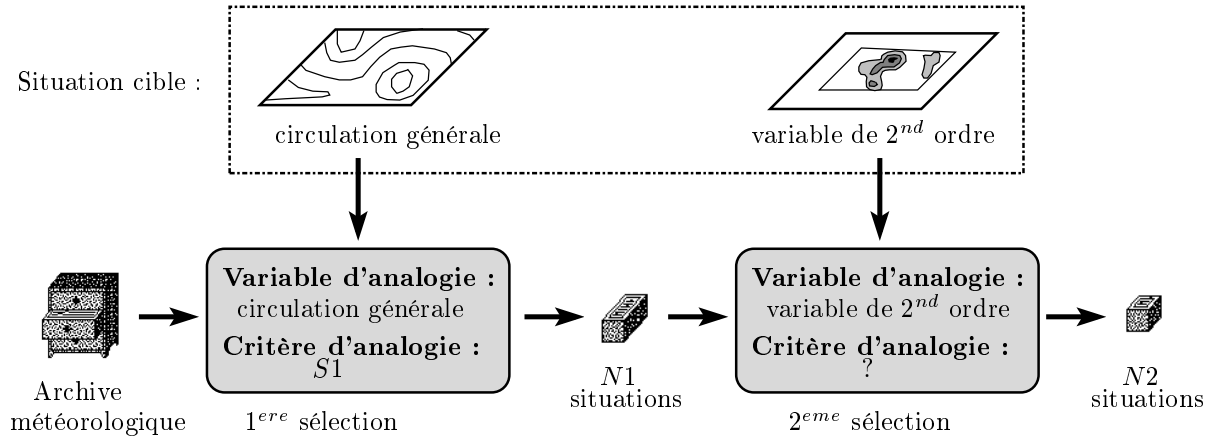


FIG. 6.1: Schéma de la sélection en deux temps des situations analogues.

Pour appliquer une recherche d'analogues en deux temps, il est donc nécessaire de choisir la variable de second ordre la plus informative, de sélectionner le critère et la fenêtre d'analogie adéquats, mais aussi d'optimiser les choix de $N1$ et $N2$. Nous voici encore une fois dans un problème combinatoire ! Notre démarche, qui consiste à nouveau à faire varier les différents paramètres un à un, les autres restant fixés de façon plus ou moins arbitraire, peut donc ne pas nous conduire exactement à la solution optimale... Au moins espérons-nous ainsi nous en rapprocher de manière significative !

Durant tout ce chapitre, nous gardons fixés les paramètres de la première sélection :

- **variable d'analogie** : $z1000+12$ et $z500+24$,
- **fenêtre d'analogie** : fenêtres statiques optimales définies sur les 7 regroupements, \textcircled{R}_1 à \textcircled{R}_7 , de nos groupements pluviométriques (cf. 5.1.4.2),
- **critère d'analogie** : $S1$.

D'après les résultats précédents, ces choix constituent un bon compromis entre performance de l'adaptation par analogie et complexité des algorithmes.

6.1.2 Tri préliminaire

En premier lieu, il est nécessaire de faire un tri au milieu de l'ensemble des nouvelles variables disponibles au sein de l'archive LTHER, issues des réanalyses NCEP/NCAR. Toutes ne sont effectivement pas aussi informatives vis à vis des précipitations, comme nous l'avons vu lors du choix de la variable d'analogie de premier ordre (cf. 4.1.2). En outre, certaines risquent d'être redondantes avec l'information déjà apportée par la circulation synoptique.

Nous les avons donc testées une par une en tant que potentielles variables de second ordre. De façon similaire au tri préliminaire effectué pour choisir la variable d'analogie de premier ordre, différents critères d'analogie ont été successivement considérés : $S1$, r et $RMSE$. Deux types de fenêtres d'analogies de second ordre ont par ailleurs été pris en compte :

- une fenêtre unique et de grande échelle (fenêtre optimale pour 2 géopotentiels, $25 \times 10^\circ$) : figure 6.2a,

- des fenêtres à petites échelles ($2.5 \times 2.5^\circ$, soit 4 points de grille), propres à chaque regroupement et centrées sur chacun d'eux : figure 6.2b.

Les résultats affichés correspondent au gain d'utilité **Comp. CRPS**, obtenu en moyenne sur l'ensemble des groupements, grâce aux $N2$ analogues issues de la seconde sélection par rapport à l'utilité des $N1$ analogues retenues lors de la première sélection. $N1$ et $N2$ ont été arbitrairement fixés à respectivement 60 et 30. Enfin, pour chaque variable, nous nous sommes placés à l'altitude et à l'heure de mesure qui produit les meilleures performances.

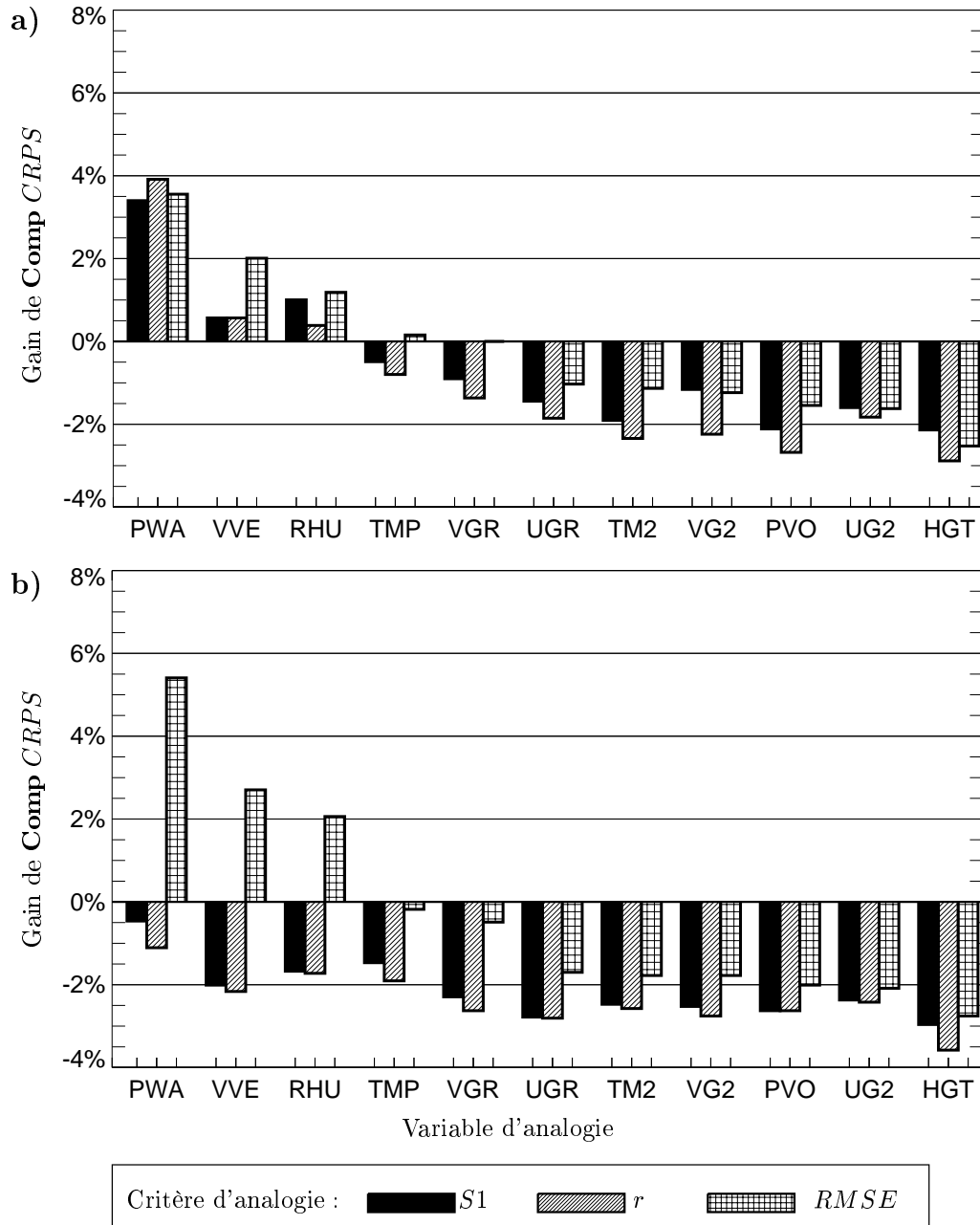


FIG. 6.2: Comparaison des gains moyens de performance sur l'ensemble des groupements, obtenus avec l'utilisation de chacune des variables de l'archive LTRE en tant que variable d'analogie de second ordre. Fenêtre d'analogie de second ordre : **a)** unique de grande échelle, **b)** : propre à chaque regroupement et de petite échelle.

Avant toute chose, on constate que l'utilisation d'une variable de second ordre n'est pas forcément un gage d'amélioration des performances. Certaines variables entraînent effectivement une dégradation de l'utilité des prévisions. Cela prouve toutefois que le gain apporté par les variables de second ordre pertinentes n'est pas lié au simple fait de passer de 60 à 30 analogues : ces gains recouvrent une réelle liaison physique avec les précipitations.

Les variables qui présentent un intérêt lors d'une analogie de second ordre sont :

- Celles liées à la mesure de la vapeur d'eau présente dans l'air : PWA (masse d'eau précipitable) et RHU (humidité relative). Ceci n'est finalement guère étonnant, la pluie n'étant jamais que la condensation puis la précipitation de la vapeur d'eau atmosphérique.
- Ainsi qu'une mesure des mouvements verticaux dans l'atmosphère : VVE (ω , vitesse verticale en coordonnées « pression »). Cette variable, qui dépend de la paramétrisation de la convection dans le modèle NWP, permet de prendre en compte les phénomènes convectifs ou les soulèvements orographiques, qui sont très importants dans la genèse des précipitations.

Ces résultats sont cohérents avec l'étude menée par Roebber et Reuter (2002) sur la comparaison détaillée de deux situations analogues qui ont engendré des cumuls de précipitations fort différents. A l'aide de simulations numériques, ces auteurs ont pu faire varier diverses caractéristiques de ces situations. Ils ont ainsi mis en évidence, dans un contexte où les conditions synoptiques sont similaires, la sensibilité des précipitations à l'humidité des parcelles d'air ainsi qu'à la force des ascendances près de la surface. Nous retrouvons ici, dans un contexte beaucoup plus général que cette étude, à la fois l'importance de l'humidité et celle des mouvements ascendants. Nous ne disposons malheureusement au sein de l'archive LTHE que du niveau 500 hPa en ce qui concerne les vitesses verticales VVE, ce qui ne correspond pas vraiment à un niveau proche du sol. Néanmoins, la plupart des mouvements verticaux se propagent sur de grandes hauteurs ; ainsi même s'il aurait été intéressant de disposer de niveaux inférieurs, le niveau 500 hPa se révèle informatif.

L'état thermique de l'atmosphère (TMP : température), n'apporte pour sa part aucune information supplémentaire à l'issue de la première sélection. A circulation synoptique comparable, la température n'est donc pas un facteur explicatif des précipitations. Notons qu'en tant que variable de premier ordre, TMP figurait en assez bonne place, devant VVE et au même niveau que RHU. Il semble donc que toute l'information contenue dans TMP soit déjà présente dans la circulation synoptique, notamment quand on considère plusieurs niveaux géopotentiels, la température étant corrélée avec l'épaisseur de l'atmosphère entre deux niveaux de pression donnés.

Les variables reliées à la circulation générale, HGT (hauteur de géopotentiels), UGR (vitesse du vent zonal) et VGR (vitesse du vent méridien), sont redondantes avec la première sélection, basée sur les similitudes de circulation synoptique. Elles dégradent même les performances, étant donné qu'elles ne font que focaliser sur quelques éléments isolés qui, comme on l'a vu précédemment, sont toujours moins informatifs que la combinaison de géopotentiels utilisée en première analogie. Il est intéressant de noter que le niveau absolu des géopotentiels, qui est ignoré par le critère d'analogie *S1* de la première sélection, se révèle effectivement tout à fait inutile. Nous

confirmons donc encore une fois que c'est la forme des géopotentiels qui importe pour rechercher des analogies, et non pas leur altitude.

Les variables définies selon des niveaux isentropiques, TM2 (température), UG2 (vitesse du vent zonal) et VG2 (vitesse du vent méridien), présentent des résultats encore moins intéressants que leurs homologues définis selon des niveaux de géopotentiel. Néanmoins, de même que lors de la sélection préliminaire effectuée pour choisir la variable d'analogie de premier ordre, nous pensons que ceci est seulement dû au fait que ces variables ne sont pas disponibles à +12h au sein de notre archive. Leurs résultats ne sont donc pas tout à fait comparables avec ceux des autres variables.

Enfin, les performances obtenues par PV0 (tourbillon potentiel) semblent indiquer que cette variable ne présente pas d'intérêt en tant que variable d'analogie de second ordre. Cependant, les travaux de Massacand *et al.* (1998) ont montré une forte liaison entre les orages qui se produisent sur le versant sud des Alpes et des anomalies de tourbillons potentiels au niveau de la tropopause, reconnues comme étant le signe d'une intrusion d'air stratosphérique. Notre absence de résultats probants est peut-être explicable par l'inadaptation de notre méthode d'analogie à cette variable. En effet, Fehlmann *et al.* (2000) ont mis en évidence l'importance d'une représentation fine de l'anomalie de PV0, or :

- D'une part nous ne disposons ici que d'une résolution de $2.5 \times 2.5^\circ$: à quel point une valeur de tourbillon potentiel sur une telle maille est-elle pertinente ?
- D'autre part, les fenêtres d'analogies utilisées lors de ce tri préliminaire peuvent ne pas convenir, d'autant plus que les anomalies qui induisent les précipitations ne se trouvent pas à l'aplomb des zones concernées par ces pluies, contrairement à ce à quoi on peut s'attendre pour les autres variables. Néanmoins, des cartographies de pertinence réalisées pour les groupements **Toce_Nord** et **Chassezac** (figure 6.3) ne font pas apparaître de zones préférentielles quant à une éventuelle analogie de second ordre sur PV0 (mesuré ici à +00h).

Ainsi, même si elle joue indubitablement un rôle sur les précipitations lors de certaines situations météorologiques, cette variable ne semble pas informative dans un contexte général, sous la forme dont nous en disposons.

En ce qui concerne les critères d'analogie, le critère de distance moyenne *RMSE* est presque systématiquement plus efficace que les critères de co-fluctuation *S1* et *r*, quelle que soit la variable considérée. Ceci est d'autant plus marqué que l'on considère des fenêtres de plus en plus petites, la notion de répartition spatiale n'étant plus réellement significative sur seulement 4 points de grille.

Par ailleurs, les meilleures performances sont obtenues avec les petites fenêtres d'analogie centrées sur les regroupements, contrairement à ce que l'on observait lors de la sélection de premier ordre, laquelle privilégiait l'utilisation d'une fenêtre de grande échelle. C'est donc bien, ainsi que l'on s'y attendait, une analogie à caractère plus local que l'on doit adjoindre à la première analogie portant sur la circulation synoptique.

Enfin, il est important de signaler que les trois variables intéressantes, PWA, RHU et VVE, sont dans les réanalyses des variables de type B, c'est à dire assez fortement influencées par le

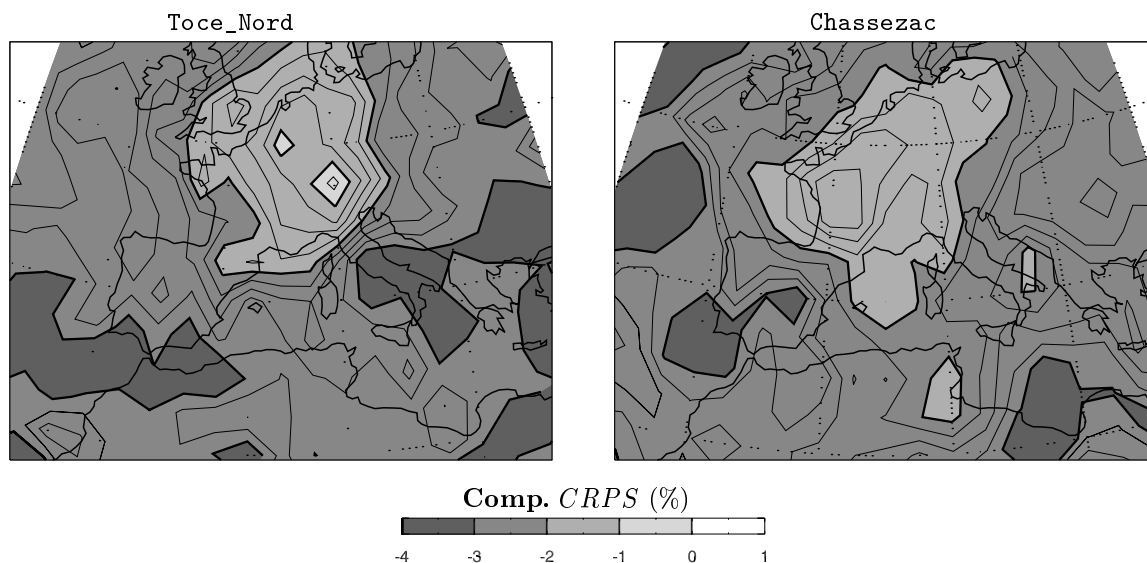


FIG. 6.3: Cartographies de pertinence pour la variable d'analogie de second ordre PV0.

modèle numérique utilisé lors de l'analyse des situations (cf. 2.1.3.2). Cela pose des questions quant à leur éventuelle utilisation dans un cadre opérationnel. Il faut en effet :

- que le modèle NWP qui prévoit la situation cible soit compatible avec celui qui a généré l'archive météorologique, afin que ces variables aient la même signification et la même résolution pour les situations candidates,
- et qu'elles soient bien prévisibles par le modèle NWP, afin que l'analogie ne porte pas sur une situation mal anticipée par le modèle opérationnel.

Outre les performances intrinsèques à chaque variable, ces considérations sont donc à prendre en compte lors du choix d'un algorithme opérationnel.

6.1.3 Définition de variables élaborées

Les variables brutes que nous avons testées ne représentent pas forcément les meilleures mesures des processus atmosphériques qui nous intéressent. Ainsi, on peut préférer exprimer la quantité de vapeur d'eau présente dans l'air sous la forme d'une humidité absolue ou d'un rapport de mélange au lieu d'une humidité relative (RHU). Il peut aussi être intéressant de considérer des variables qui correspondent à des processus physiques, tels que des flux d'humidité.

Nous allons donc construire différentes variables élaborées à partir des variables brutes dont nous disposons. Afin d'estimer l'intérêt de ces nouvelles variables, nous les testerons comme variable de second ordre avec le même protocole que celui utilisé précédemment. Cependant, au vu des résultats déjà obtenus, nous nous limiterons :

- d'une part à l'utilisation de fenêtres de petite échelle : fenêtres propres à chaque regroupement de $2.5^\circ \times 2.5^\circ$ (centrées sur chaque regroupement),
- et d'autre part à l'utilisation du critère d'analogie *RMSE*.

6.1.3.1 Expression de l'humidité

Le tri préliminaire a montré l'intérêt des variables liées à la présence de vapeur d'eau (PWA et RHU) en tant que variable de second ordre. Néanmoins, elles ne constituent peut-être pas ici la meilleure façon de mesurer l'état hygrométrique des masses d'air.

En effet, l'humidité relative indique par exemple si l'air est proche ou non de son niveau de saturation en vapeur d'eau. Or, plus l'air est chaud, plus il peut en contenir une quantité importante. RHU est donc un paramètre important, mais qui ne nous renseigne pas sur la quantité totale de vapeur d'eau présente dans l'air. Nous avons donc calculé l'humidité spécifique q , qui représente la masse de vapeur d'eau présente dans une masse d'air unité. Elle correspond donc à la quantité absolue d'humidité présente dans l'air à un niveau donné. Elle est généralement exprimée en g.g^{-1} et s'obtient par l'intermédiaire de la relation suivante :

$$q = \epsilon \cdot \frac{e_{sat} \cdot RHU}{P + (\epsilon - 1) \cdot e_{sat} \cdot RHU} \quad (6.1)$$

où :

- ϵ est le rapport des masses molaires de l'air sec et de la vapeur d'eau : $\epsilon = 0.622$,
- e_{sat} est la pression de vapeur saturante,
- P est la pression atmosphérique.

La pression de vapeur saturante dépend de la température de l'air TMP selon l'équation de Clausius-Clapeyron. Il existe de nombreuses formulations approchées de l'expression de la pression de vapeur saturante ; nous avons choisi d'utiliser ici la relation de Murray (1967), qui s'exprime de la manière suivante :

$$e_{sat}(TMP) = 6.1078 \times \exp \frac{a \times (TMP - 273.15)}{TMP - b} \quad (6.2)$$

où TMP s'exprime en degrés Kelvin. Les coefficients a et b dépendent de l'environnement : eau liquide (condensation liquide) ou glace (condensation solide), comme indiqué dans le tableau 6.1.

	eau liquide	glace
a	17.2693882	21.8745584
b	35.86	7.66

TAB. 6.1: Paramètres intervenant dans la formulation de la pression de vapeur saturante.

L'eau précipitable PWA, disponible dans l'archive LTHE, correspond à l'intégrale sur toute la colonne d'air de q . Ces deux variables sont donc fortement liées entre elles.

Nous avons par ailleurs créé deux variables supplémentaires, en considérant que pour que des précipitations importantes aient lieu, il est préférable d'avoir à la fois une forte humidité absolue (présence de beaucoup de vapeur d'eau), et une importante humidité relative (condensation plus rapide de la vapeur d'eau). Nous définissons ainsi PW% (resp. $q\%$) comme le produit de l'eau précipitable PWA (resp. l'humidité absolue q) par l'humidité relative :

$$PW\% = PWA \times RHU \quad (6.3)$$

$$q\% = q \times RHU \quad (6.4)$$

Ces variables ne recouvrent aucune grandeur physique *réelle* : elles correspondent plutôt à une schématisation d'un processus physique.

Nous disposons ainsi de 5 expressions différentes de l'humidité : PWA, RHU, q , PW% et $q\%$. La figure 6.4 présente les performances obtenues grâce à leur utilisation en tant que variables de second ordre. On constate ainsi que les variables PW% et $q\%$, bien que très artificielles, sont plus informatives que les variables physiques sur lesquelles elles se basent : PWA et q . Les mesures absolues et relatives de l'humidité sont donc bien complémentaires pour expliquer les précipitations.

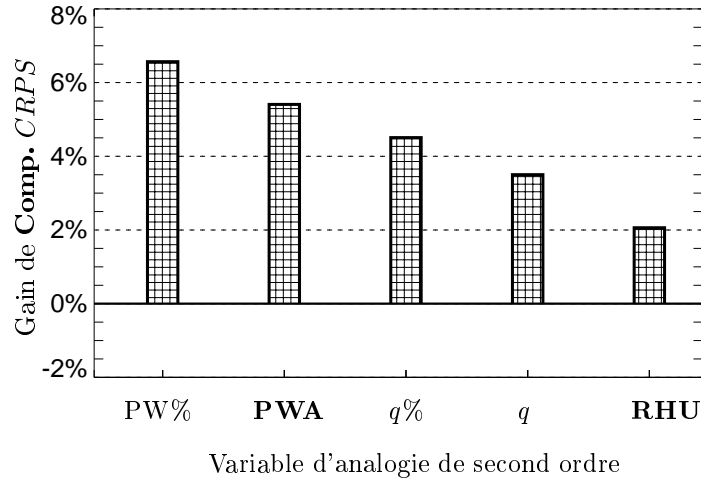


FIG. 6.4: Comparaison des gains moyens de performance sur l'ensemble des groupements, obtenus avec l'utilisation de variables d'humidité en tant que variables d'analogie de second ordre.

Il apparaît par ailleurs qu'il est préférable d'exprimer l'humidité absolue comme intégrée sur toute la colonne d'air (PW%, PWA), plutôt que mesurée au seul niveau 850 hPa ($q\%$, q). Toutefois, l'humidité atmosphérique se concentre principalement entre les niveaux sol et 500 hPa. Ne disposant pas de données d'humidité pour des niveaux inférieurs à 850 hPa, il ne nous est donc pas possible de savoir si l'intérêt de PWA et de PW% réside dans le fait que ces variables reflètent l'humidité des très basses couches, primordiale dans la genèse des précipitations, ou s'il est réellement dû à la prise en compte d'un cumul d'humidité plutôt qu'à une mesure ponctuelle.

6.1.3.2 Flux horizontaux d'humidité

Plus que la présence d'une forte humidité, ce qui rend possible de forts cumuls de précipitations est souvent le fait qu'une cellule précipitante puisse être alimentée en humidité, comme cela était par exemple le cas lors des pluies catastrophiques survenues les 8 et 9 septembre 2002 dans la région du Gard (Delrieu et al. 2003, cf. chapitre 9). Nous avons donc construit un flux d'humidité spécifique $V \times q$ en calculant le produit de la vitesse du vent et de l'humidité présente :

$$V \times q = q \cdot \sqrt{VGR^2 + UGR^2} \quad (6.5)$$

où : $\left| \begin{array}{l} VGR \text{ est la vitesse du vent méridien,} \\ UGR \text{ est la vitesse du vent zonal.} \end{array} \right.$

Par analogie avec le flux d'humidité spécifique, nous définissons les variables $V \times PW$ et $V \times PW\%$ comme les produits de la vitesse du vent et de l'eau précipitable :

$$V \times PW = PWA \cdot \sqrt{VGR^2 + UGR^2} \quad (6.6)$$

$$V \times PW\% = PW\% \cdot \sqrt{VGR^2 + UGR^2} \quad (6.7)$$

Bien que pouvant y être assimilées, ces variables ne constituent pas de réels flux :

- La vitesse du vent est ici considérée à un seul niveau (850 hPa) alors que PWA intègre tous les niveaux possibles. Pour être parfaitement rigoureux, il aurait fallu intégrer les flux d'humidité sur toute la colonne d'air, mais nous ne disposons malheureusement pas des données nécessaires pour le faire. Néanmoins, $V \times PW$ peut être perçu comme une approximation du flux d'eau précipitable, celle-ci se concentrant surtout dans les basses couches.
- $PW\%$ ne correspond pas à une grandeur physique réelle : il faut là aussi voir $V \times PW\%$ comme une schématisation d'un processus physique.

On note par ailleurs qu'un grand nombre de groupements de l'archive EDF est situé dans la vallée du Rhône, qui présente une direction Nord-Sud privilégiée. Il est ainsi apparu, lors du tri préliminaire, que les vents méridiens ou zonaux n'étaient pas aussi informatifs l'un que l'autre vis à vis des précipitations. Il peut donc être intéressant de différencier les flux d'humidité venant selon un axe Nord-Sud de ceux venant selon un axe Est-Ouest. C'est pourquoi nous avons aussi calculé les flux directionnels méridiens et zonaux associés aux trois flux d'humidité (ou assimilés comme tel) $V \times q$, $V \times PW$ et $V \times PW\%$, en adoptant les notations suivantes :

$$V \times X_{NS} = X \times VGR \quad (6.8)$$

$$V \times X_{EO} = X \times UGR \quad (6.9)$$

où : $\left| \begin{array}{l} V \times X_{NS} \text{ est le flux méridien de la variable } X \text{ (direction Nord-Sud),} \\ V \times X_{EO} \text{ est le flux zonal de la variable } X \text{ (direction Est-Ouest).} \end{array} \right.$

Les performances obtenues par ces différents flux horizontaux d'humidité, utilisés comme variables d'analogie de second niveau (figure 6.5), se révèlent en fait moins intéressantes que celles permises par la seule utilisation des variables d'humidité. Ces résultats peuvent paraître étonnants au premier abord, les flux d'humidités ayant une relation physique plus forte avec les précipitations. Il faut cependant garder à l'esprit que l'analogie de second ordre intervient après une première analogie sur la circulation synoptique. Les situations analogues présentent donc des vitesses moyennes de vent à peu près similaires. Les différences entre les situations analogues sont

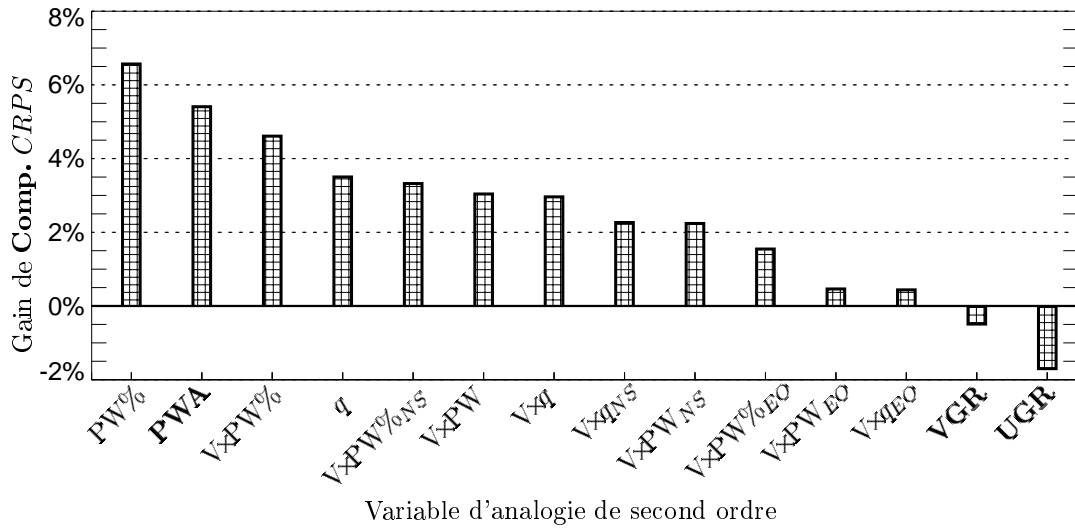


FIG. 6.5: Comparaison des gains moyens de performance sur l'ensemble des groupements, obtenus avec l'utilisation de flux horizontaux d'humidité en tant que variables d'analogie de second ordre.

donc plus marquées en terme d'humidité qu'en terme de flux d'humidité, ce qui peut expliquer l'utilité moindre de ces derniers.

La discrimination des flux horizontaux d'humidité selon leur direction ne semble pas présenter un grand intérêt : les performances associées sont toujours inférieures à celles des flux horizontaux totaux. On remarque toutefois que les composantes zonales apportent des gains de performance quasi-nuls et que la plus grande part de l'information est contenue dans la composante méridienne des flux. Il existe donc bien une composante privilégiée, sûrement induite par la répartition géographique des groupements utilisés.

Flux verticaux d'humidité

La vitesse verticale des particules d'air (VVE) est la seconde variable brute la plus informative, juste après PWA. Elle joue en effet un rôle important dans la production des précipitations, l'ascendance d'une masse d'air humide étant l'un des principaux processus générateur. C'est pourquoi nous avons introduit un pseudo¹ flux vertical d'humidité absolue $\omega \times q$:

$$\omega \times q = q \times VVE \quad (6.10)$$

De même que précédemment, nous construisons par analogie des variables similaires, basées sur l'utilisation de l'eau précipitable et de PW% :

¹Nous employons ici le terme *pseudo* car la vitesse verticale étant exprimée en coordonnées pression, la variable que nous créons ainsi n'a pas vraiment la dimension d'un flux : VVE s'exprime ici en Pa.s^{-1} . A titre d'ordre de grandeur, une variation verticale de 1 Pa correspond à peu près :

- dans les basses couches de l'atmosphère, à une distance de 0.1 m,
- à un niveau de 500 hPa (niveau de mesure de VVE), à 0.15 m.

$$\omega \times PW = PWA \times VVE \quad (6.11)$$

$$\omega \times PW\% = PW\% \times VVE \quad (6.12)$$

Comme le montre la figure 6.6, les flux d'humidité verticaux sont plus informatifs que les flux horizontaux. Ils indiquent que les précipitations, à circulation synoptique donnée, sont en grande partie conditionnées à la fois par la présence d'humidité et par la possibilité pour celle-ci de se condenser via une ascendance (forçage orographique, convection thermique, ...). Ce principe est bien sûr connu depuis longtemps, mais il est intéressant de voir que son impact peut au moins partiellement être résumé via ces simples pseudo flux verticaux d'humidité.

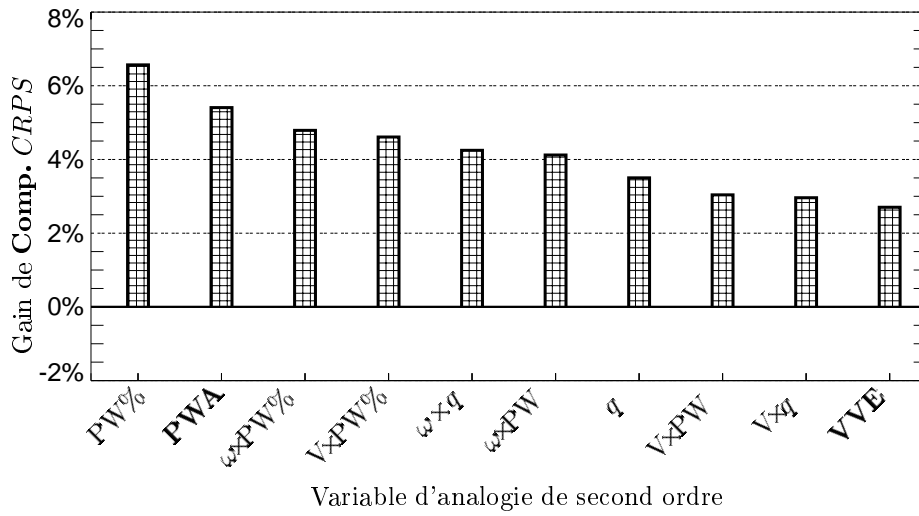


FIG. 6.6: Comparaison des gains moyens de performance sur l'ensemble des groupements, obtenus avec l'utilisation de flux verticaux d'humidité en tant que variables d'analogie de second ordre.

Néanmoins, sauf en ce qui concerne $\omega \times q$, ils restent moins performants que les seules variables d'humidité. PWA et PW% étant des constructions relativement artificielles, les flux associés peuvent ne plus revêtir autant de sens physique qu'attendu.

6.2 PW% comme variable de second ordre

Lorsque l'on considère tous les groupements pêle-mêle, la variable d'analogie de second ordre la plus pertinente s'avère être PW%, c'est à dire le produit de l'humidité relative (RHU) mesurée à 850 hPa et de l'eau précipitable (PWA). Nous allons donc nous pencher dans cette section sur l'optimisation de l'algorithme d'analogie en considérant uniquement cette variable de second ordre.

Nous conserverons ici le cadre de la performance globale sur l'ensemble des groupements. Les résultats ainsi obtenus nous serviront, dans la section suivante, de point de départ pour appréhender les spécificités de chaque groupement vis à vis de la variable d'analogie de second ordre.

6.2.1 Choix de l'heure de mesure

Dans un premier temps, nous étudions l'évolution des gains de performance en fonction de l'heure de mesure de PW% (tableau 6.2). Ces résultats sont obtenus en considérant comme précédemment des fenêtres de second niveau de $2.5 \times 2.5^\circ$ propres à chaque regroupement (\textcircled{R}_1 à \textcircled{R}_7) et en conservant $N1$ (resp. $N2$) fixé à 60 (resp. 30).

	Heure de mesure (hTU)					
	-12	+00	+12	+24	+36	+48
Gain de compétence (%)	- 1.03	+ 0.93	+ 6.57	+ 6.10	+ 1.75	- 1.42

TAB. 6.2: Evolution des gains de compétence (**Comp.** *CRPS*) moyens sur l'ensemble des groupements en fonction du choix de l'heure de mesure de PW%.

Les mesures à +12h et +24h apparaissent comme étant les plus performantes. De même que pour les variables de premier ordre, l'heure de mesure de la variable de second niveau doit donc être compatible avec la période de cumul du prédictand (de +6h à +30h). Si l'on s'en écarte, on assiste alors à des pertes de performances significatives.

La variable de second niveau doit donc elle aussi être prévue par un modèle NWP : la simple observation de PW% n'apporte aucune aide au delà de 24 heures. Ceci implique que les modèles NWP doivent être capables de produire des prévisions de bonne qualité de PWA et de RHU. Cependant, même dans le cas contraire, le travail présenté dans ce chapitre n'est pas inutile pour autant. Il peut en effet être considéré comme une étude diagnostique de la dépendance des précipitations à l'analogie sur les variables d'humidité.

Etant donnée l'importance que revêt l'heure de mesure, il peut être intéressant d'utiliser ici aussi, comme pour la première sélection, une combinaison de plusieurs heures de mesures. Pour ce faire, afin de classer entre elles les situations candidates J_{ana} par rapport à la journée cible C , nous utilisons à nouveau une distance d'analogie DA basée sur la moyenne arithmétique des critères d'analogie $RMSE$ calculés sur chacun des champs de la combinaison :

$$DA(J_{ana}) = \frac{1}{N_{Ch}} \sum_{i_{Ch}=1}^{N_{Ch}} RMSE \left(PW\%_{i_{Ch}}^C; PW\%_{i_{Ch}}^{J_{ana}} \right) \quad (6.13)$$

où :

- N_{Ch} est le nombre de champs de PW% intervenant comme variable d'analogie de second niveau,
- $PW\%_i^J$ est le $i^{\text{ème}}$ champ de PW% des N_{Ch} champs intervenant comme variable d'analogie, de la situation J .

Le tableau 6.3 présente les résultats obtenus pour quelques combinaisons. On constate ici aussi que :

- il est utile de combiner les échéances +12h et +24h,
- et la prise en compte d'échéances situées en dehors de la période de cumul des précipitations n'apporte aucun bénéfice.

	Heure de mesure (hTU)			
	+00 & +24	+00 & +12	+12 & +24	+00 & +12 & +24
Gain de compétence (%)	+ 5.72	+ 5.51	+ 8.73	+ 7.75

TAB. 6.3: Evolution des gains de compétence (**Comp. CRPS**) moyens sur l'ensemble des groupements pour différentes combinaisons de champs PW%.

6.2.2 La fenêtre d'analogie de second ordre

Nous allons maintenant nous intéresser au choix de la fenêtre de second ordre. Dans toute cette section, la variable d'analogie sur laquelle se base la seconde sélection est celle qui a été reconnue comme optimale jusqu'ici : PW% mesurée à +12h & +24h. En outre, nous continuons d'adopter $N1 = 60$ et $N2 = 30$.

6.2.2.1 Cartographie de pertinence

En premier lieu, nous reprenons le principe des cartographies de pertinence exposé dans la partie 5.1.1.1, en l'appliquant à l'analogie de second ordre. Deux modifications sont toutefois apportées par rapport aux utilisations précédentes :

- la pertinence est exprimée sous forme de gain de performance (**Comp. CRPS**) par rapport à la sélection de premier ordre,
- la *fenêtre élémentaire* se réduit à un unique point de grille (ceci est rendu possible par l'utilisation du critère d'analogie *RMSE*).

Ces cartographies permettent d'appréhender où est localisée l'information supplémentaire que peut apporter la variable de second ordre, en plus de celle déjà contenue dans la première sélection.

La figure 6.7a présente la cartographie de pertinence obtenue en considérant les gains de performances moyens sur l'ensemble des groupements EDF. La première chose que l'on remarque, c'est que la zone pertinente est bien plus restreinte ($\approx 8^\circ \times 8^\circ$) que celle de l'analogie de premier ordre ($\approx 30^\circ \times 20^\circ$, cf. 5.1.1). On assiste même à des pertes de performances dès que l'on sort de l'aire de répartition des groupements. En outre, la zone pertinente est centrée sur le barycentre des groupements et épouse leur répartition géographique (en forme de L couché sur le dos).

L'analogie de second ordre apparaît ainsi comme une analogie assez locale : elle doit être considérée dans les environs immédiats du groupement cible. Nous ne nous plaçons donc plus maintenant aux échelles synoptiques.

Ce caractère local est confirmé par les figures 6.7b et 6.7c qui présentent les cartographies de pertinence associées respectivement aux groupements **Hte_Durance** et **Pyrénées_Atl._Ht_Bassin**. Les zones pertinentes se révèlent ici aussi de petite taille et centrées sur les groupements cibles. On remarque par ailleurs, notamment en ce qui concerne le groupement **Hte_Durance**, que la répartition de la pertinence n'est pas forcément isotrope autour du groupement : elle s'organise ici selon un axe privilégié SO-NE. Cette organisation semble être le reflet des principaux flux générateurs de précipitations sur ce groupement : flux de Sud et retour d'Est.

Lorsque l'on considère les regroupements au sein desquels s'insèrent ces groupements (figures 6.7d et 6.7e, respectivement \textcircled{R}_7 et \textcircled{R}_2), on observe que les zones pertinentes restent sensiblement les mêmes et ne s'étendent pas. Les regroupements, tels que nous les avons effectués, semblent donc être relativement adaptés à l'échelle locale de la variable de second ordre.

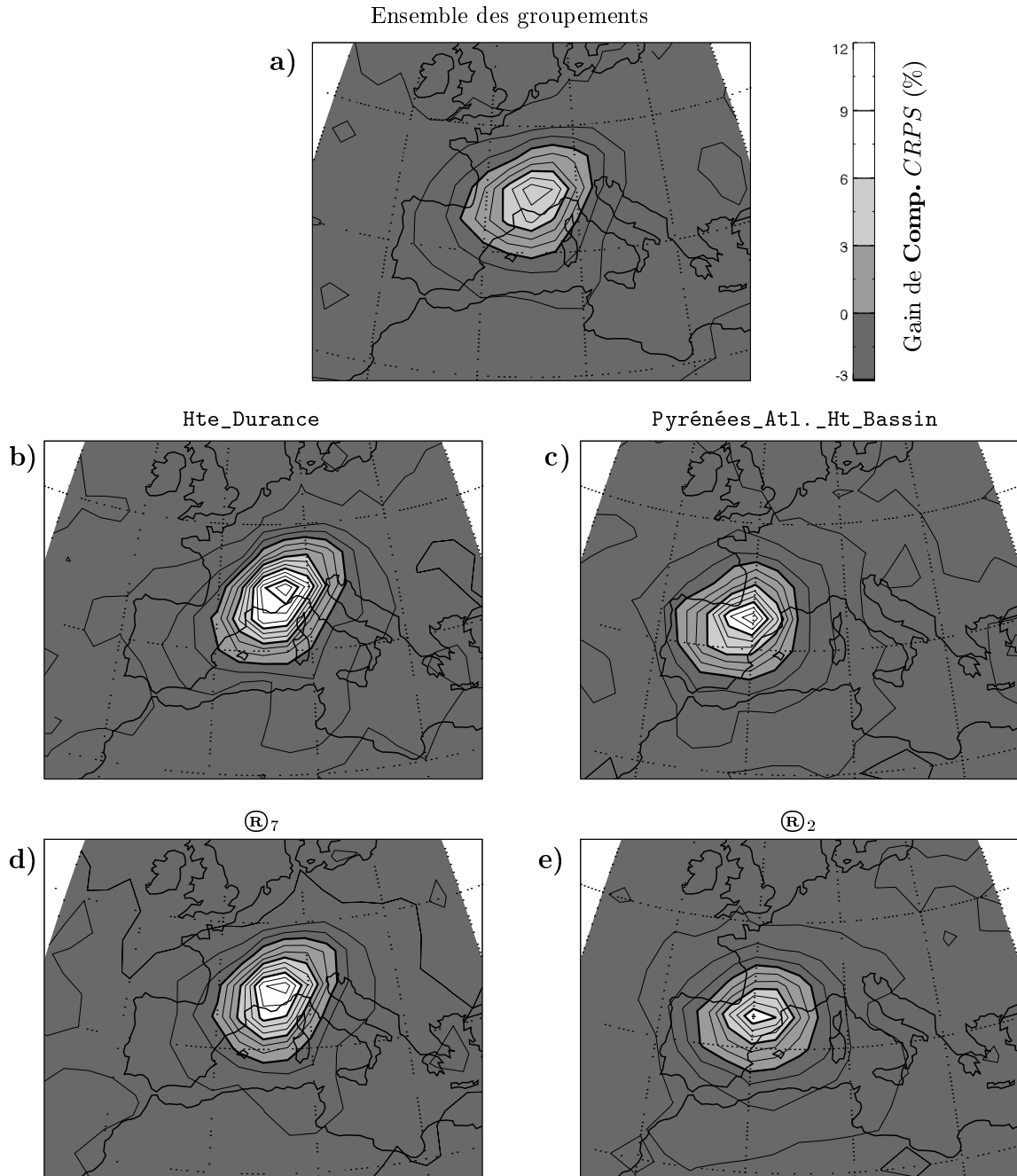


FIG. 6.7: Cartographie de pertinence exprimée en gain de performance (**Comp. CRPS**) par rapport à la première sélection. Variable de second ordre : $PW\%$ mesuré à +12h & +24h. **a)** ensemble des groupements de l'archive EDF; **b)** groupement Hte_Durance; **c)** groupement Pyrénées_Alt._Ht_Bassin; **d)** regroupement \textcircled{R}_7 ; **e)** regroupement \textcircled{R}_2 .

En s'intéressant à des variables locales, on risque cependant de se retrouver confronté à l'une des limitations des réanalyses NCEP/NCAR : leur faible résolution spatiale de $2.5^\circ \times 2.5^\circ$. Avec une telle résolution, les zones réellement pertinentes ne sont décrites que par 4 points de grille. Cela peut s'avérer être un échantillonnage insuffisant pour capter le maximum d'information. En outre, les données que l'on manipule représentent plus ou moins des valeurs moyennes sur une maille du modèle. Elles cachent donc toute la variabilité de ces variables météorologiques aux échelles inférieures à 2.5° . Tant que l'on s'intéressait à des champs de géopotentiels, cette résolution était adaptée à la variable considérée, mais maintenant que l'on s'attache à des variables de portée¹ plus faible et dont les zones pertinentes sont très restreintes, il serait sûrement souhaitable de disposer d'une plus grande résolution.

6.2.2.2 Optimisation de la fenêtre

Les cartographies que nous venons de réaliser ont montré que même si les zones qui apportent réellement de l'information sont d'extensions limitées, la répartition de la pertinence demeure relativement lisse. Nous allons donc nous limiter à la recherche de fenêtres optimales rectangulaires, en utilisant le même algorithme d'optimisation que lors de la recherche de la fenêtre optimale de premier ordre (cf. 5.1.2). Toutefois, étant donné le caractère local de la variable de second ordre, il est impératif d'apporter une attention toute particulière au choix de la fenêtre initiale : nous sélectionnons pour cela le point de grille présentant la plus forte pertinence pour l'ensemble des groupements considérés (dénommé *point caractéristique* dans le chapitre précédent).

La figure 6.8, qui représente les résultats de l'optimisation de la fenêtre de second ordre en considérant les performances moyennes sur l'ensemble des groupements, est à comparer avec la figure 5.6 qui concernait l'analogie de premier ordre. La dimension diagonale optimale est ici de 1000 km, alors qu'elle était plutôt de l'ordre de 2000 km pour les fenêtres de premier ordre.

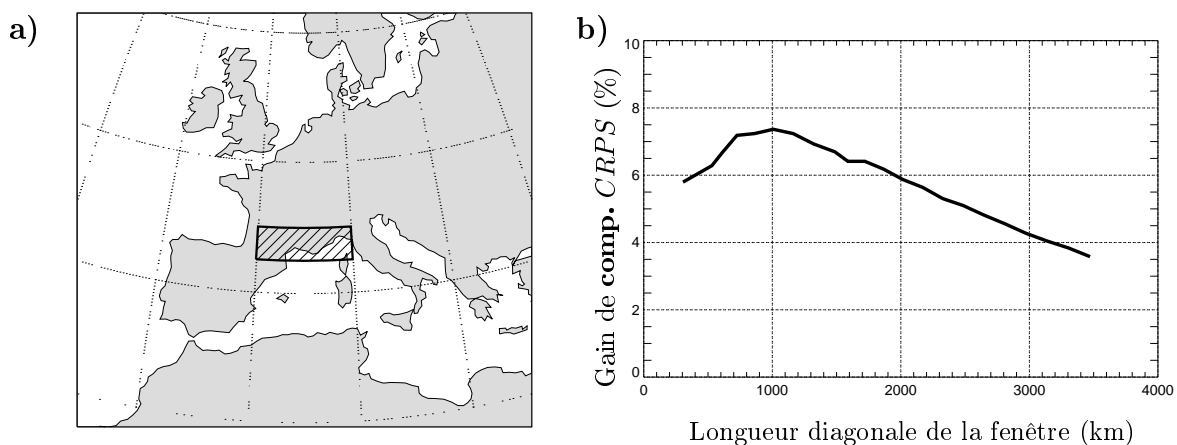


FIG. 6.8: Recherche de la fenêtre de second ordre optimale **a)** fenêtre optimale pour PW% +12h & +24h ; **b)** évolution des gains de performances (**Comp. CRPS**) en fonction de la longueur diagonale des différentes fenêtres étapes.

¹La portée correspond à la distance moyenne au delà de laquelle deux mesures d'une même variable sont décorréliées.

Nous avons par ailleurs recherché les fenêtres optimales pour chacun des 7 regroupements établis dans le chapitre précédent. Ces résultats, présentés par le tableau 6.4, montrent que ces fenêtres sont de taille encore inférieure à la fenêtre optimale pour l'ensemble des groupements. Elles se limitent en général à un domaine de $2.5^\circ \times 2.5^\circ$, soit à peine 4 points de grille. Seul le regroupement \textcircled{R}_1 , qui concerne uniquement le groupement **Bretagne**, présente un comportement à part, avec une fenêtre optimale de relativement grande taille.

	Paramètres de la fenêtre optimale			
	lon0 ($^\circ\text{E}$)	lat0 ($^\circ\text{N}$)	dlon ($^\circ$)	dlat ($^\circ$)
Fenêtre commune	0	42.5	10	2.5
Fenêtre propre à \textcircled{R}_1	-7.5	45	7.5	5
Fenêtre propre à \textcircled{R}_2	-2.5	42.5	5	0
Fenêtre propre à \textcircled{R}_3	7.5	40	2.5	2.5
Fenêtre propre à \textcircled{R}_4	5	45	5	0
Fenêtre propre à \textcircled{R}_5	5	45	2.5	2.5
Fenêtre propre à \textcircled{R}_6	2.5	42.5	2.5	2.5
Fenêtre propre à \textcircled{R}_7	5	42.5	2.5	2.5

TAB. 6.4: Fenêtres d'analogie de second ordre optimales en fonctions des regroupements considérés.

En utilisant la fenêtre commune optimisée, nous obtenons un score d'utilité, en compétence du *CRPS*, de 41.69 %, soit 7.37 % de gain par rapport à la première analogie. Grâce à l'utilisation de fenêtres propres aux regroupements, ce score passe à 42.35 % (soit 9.07 % de gain). La différence entre ces deux types de fenêtre reste donc relativement faible, mais néanmoins significative.

Toutefois, si l'on considère les différences de performance groupement par groupement (figure 6.9), on constate qu'elles sont plus marquées pour les groupements excentrés. Ainsi, la fenêtre commune induit une perte de performance sur le groupement **Bretagne**, ce que l'on peut éviter en considérant des fenêtres propres à chaque regroupement. La fenêtre commune est une fenêtre finalement assez pertinente pour le barycentre global des groupements : tous ceux qui sont situés dans cette zone ne voient qu'une légère amélioration de leur performance du fait de l'utilisation d'une fenêtre propre. Les groupements de l'archive EDF étant relativement concentrés, le gain de performance global est alors peu sensible, tandis que ponctuellement, il peut ne pas être négligeable. C'est pourquoi, pour éviter de pénaliser les groupements excentrés, l'emploi de fenêtres par regroupement est préconisé.

Cependant, il n'est pas utile d'optimiser rigoureusement ces fenêtres par regroupements. Il est en effet apparu que leur optimisation conduit quasi systématiquement à l'adoption de la fenêtre élémentaire de 4 points de grilles ($2.5^\circ \times 2.5^\circ$) centrée sur le regroupement. Si l'on utilise systématiquement de telles fenêtres, les performances obtenues sont équivalentes à celles des fenêtres optimisées, aussi bien sur l'ensemble des groupements (**Comp.** *CRPS* = 42.22 %) que sur chacun d'eux considéré individuellement.

On remarque que deux groupements de \textcircled{R}_2 (**Ht_Agout-Arn-Dourdou** et **Aveyron-Tarn_If-Garon_If**) ainsi que deux groupements de \textcircled{R}_6 (**Vezere-Vienne-Dordogne** et **Creuse-Cher-Sioule**) présentent de meilleures performances avec la fenêtre commune plutôt qu'avec des fenêtres optimisées sur leur regroupements respectifs. Ces deux regroupements ne sont donc pas

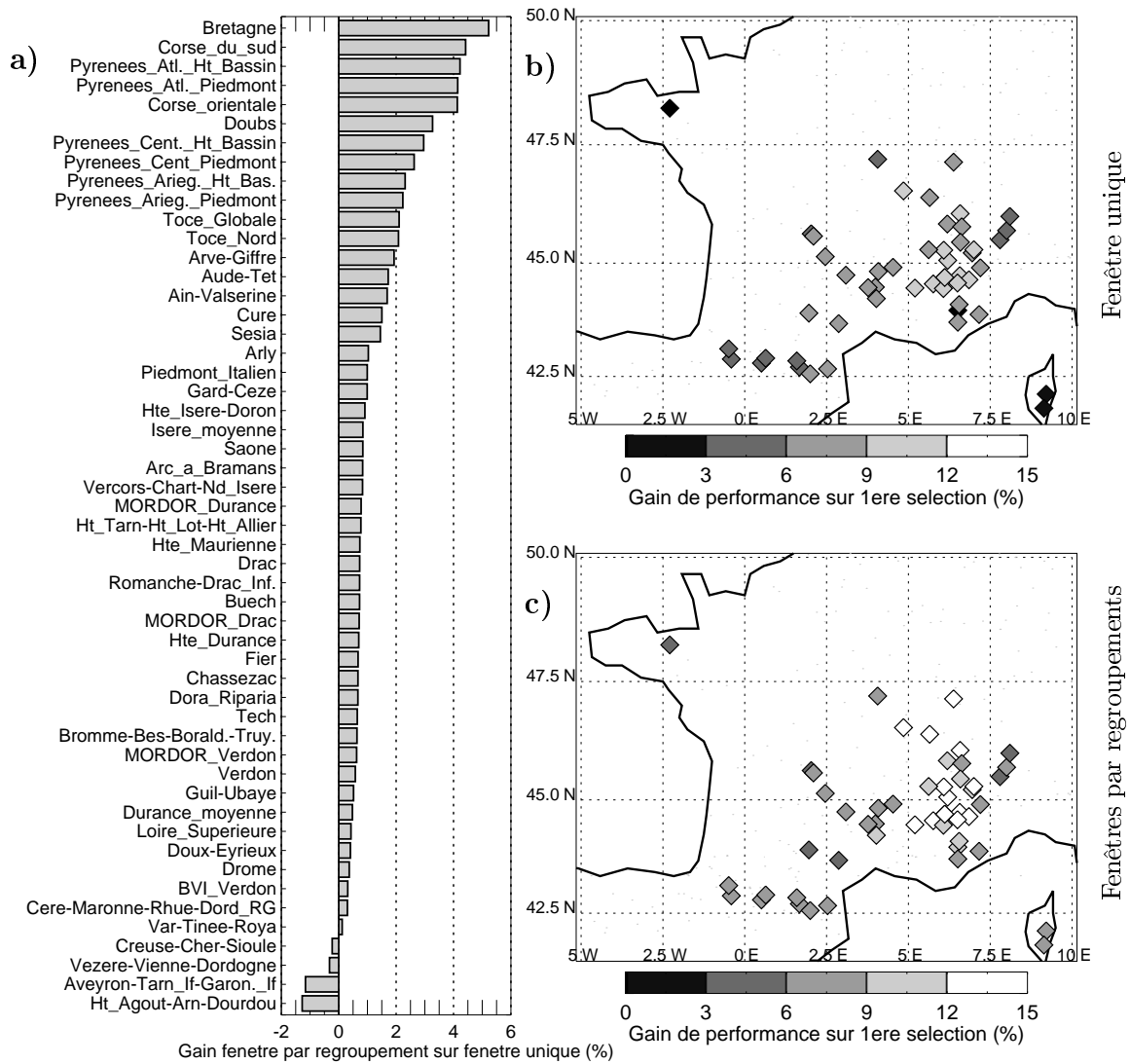


FIG. 6.9: Comparaison des performances obtenues en utilisant des fenêtres de second ordre communes à l'ensemble des groupements, ou propres à chaque regroupement. **a)** Gain de performance des fenêtres propres à chaque regroupement par rapport à la fenêtre commune; **b)** Gain de performance de la fenêtre commune par rapport à la 1^{ère} sélection; **c)** Gain de performance des fenêtres propres à chaque regroupement par rapport à la 1^{ère} sélection. (les performances sont exprimées en **Comp. CRPS**).

vraiment homogènes au sens de l'analogie de second ordre. Celle-ci ayant un caractère beaucoup plus local que la circulation synoptique, certains regroupements pourraient mériter d'être plus fins.

6.2.3 Choix de $N1$ et $N2$

Il nous reste enfin à optimiser les paramètres $N1$ et $N2$, qui concernent le nombre de situations reconnues comme analogues à l'issue de chacune des deux sélections successives. Pour étudier l'influence de ces paramètres, nous adoptons les choix résultant des optimisations précédentes :

- **variable d'analogie de second ordre** : PW% mesuré à +12h & +24h,

- **critère d'analogie de second ordre** : $RMSE$,
- **fenêtre d'analogie de second ordre** : fenêtre de $2.5^\circ \times 2.5^\circ$ centrée sur chaque regroupements \textcircled{R}_1 à \textcircled{R}_7 .

Les paramètres $N1$ et $N2$ ne sont bien sûr pas indépendants. En effet :

- Si $N1$ est petit, alors toutes les situations retenues sont très ressemblantes au sens de la circulation synoptique. Par contre, on risque de disposer de peu de situations fortement analogues au niveau de la variable de second ordre.
- Si $N1$ est grand, alors nous disposons d'un nombre important de situations analogues au niveau de la variable de second ordre, mais dont malheureusement beaucoup présentent une faible similitude au sens de la circulation générale.

Une évolution des performances obtenues en fonction des choix de $N1$ et $N2$ est présentée par la figure 6.10. On constate en premier lieu que si l'on se limite à la première sélection, la valeur optimale de $N1$ se situe autour de 30 à 35, pour un score maximal de **Comp. CRPS** = 39.24 %. Si le nombre d'analogues retenues est inférieur, les performances chutent rapidement, jusqu'à devenir négatives pour $N1 = 1$. Ce comportement extrême est cependant dû en grande partie au choix de notre mesure de l'utilité, laquelle pénalise fortement les prévisions déterministes dès qu'elles ne sont pas rigoureusement exactes. Il n'en demeure pas moins qu'il est préférable de considérer un nombre conséquent d'analogues, ce qui permet de gagner en Justesse de la prévision.

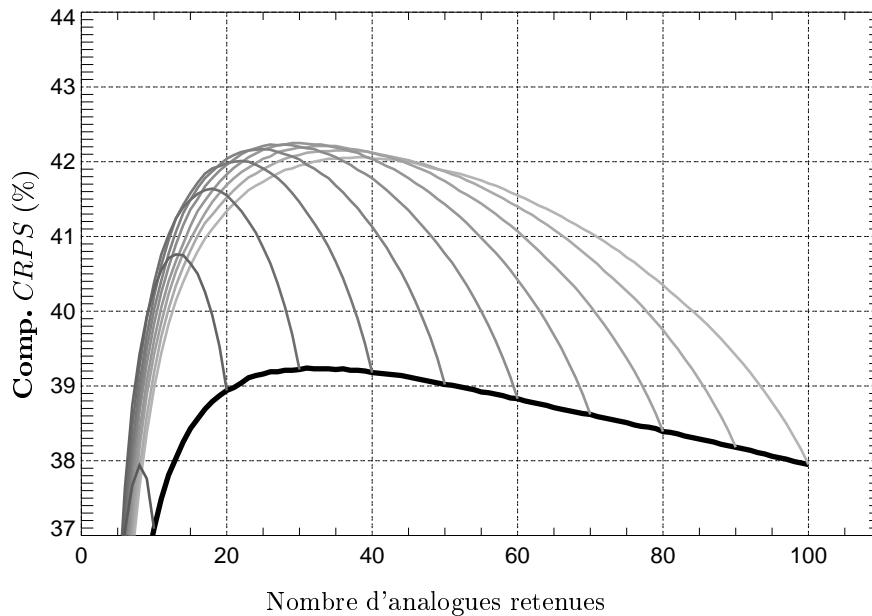


FIG. 6.10: Evolution des performances (**Comp. CRPS**) en fonction de $N1$ et $N2$. **Courbe noire** : performances à l'issue de la première sélection en fonction de $N1$; **courbes grises** : performances à l'issue de la seconde sélection, en fonction de $N2$ (chaque courbe correspond à un choix différent de $N1$; pour $N2 = N1$, les courbes grises rejoignent la courbe noire).

Si l'on s'intéresse maintenant à la seconde sélection, il apparaît que, quelles que soient les valeurs $N1$ et $N2$ choisies, les performances sont améliorées à l'issue de la seconde sélection (les

courbes grisées se trouvent au dessus de la courbe noire). Pour une valeur $N1$ fixée, la valeur optimale de $N2$ se situe aux alentours de $\frac{N1}{2}$: un peu moins si $N1$ est élevé (plus on retient d'analogues, plus on risque de considérer des situations peu pertinentes vis à vis de la circulation synoptique), et un peu plus si $N1$ est faible (afin de garder un échantillonnage suffisant de la fonction de répartition).

Finalement, les valeurs optimales sont de l'ordre de $N1 = 70$ et $N2 = 30$, pour une utilité de **Comp. CRPS** = 42.25 %. On remarque toutefois que ces utilités optimales sont assez peu sensibles au choix de $N1$: pour toute valeur de $N1$ comprise entre 100 et 40, elle demeure supérieure à 42%.

6.2.4 Régionalisation du choix de la variable de second ordre

Les résultats présentés jusqu'ici concernaient les performances moyennes sur l'ensemble des groupements. Néanmoins, lorsque l'on considère chaque groupement individuellement, on s'aperçoit que la hiérarchie des variables de second ordre n'est pas forcément toujours la même. A titre d'illustration, la figure 6.11 présente les gains de performances obtenus groupement par groupement pour 4 variables de second ordres différentes : PWA, PW%, VxPW%, ω xPW%.

PW% est ainsi la meilleure variable de second ordre d'un point de vue global, mais s'avère par exemple dépassée par VxPW% sur les groupements situés en Dordogne ou dans les Pyrénées Atlantiques. A l'inverse, VxPW% est particulièrement peu informatif en ce qui concerne les groupements du pourtour méditerranéen. Chaque variable de second ordre peut donc être à la fois très explicative pour certains groupements, et presque inutile pour d'autres. Ceci dit, PW% constitue un compromis parfaitement acceptable pour tous les groupements. Les gains que l'on peut obtenir en choisissant à chaque fois la variable la plus adaptée sont extrêmement faibles (**Comp. CRPS** passe de 42.22 à seulement 42.27 % en moyenne sur l'ensemble).

On remarque en outre une certaine homogénéité régionale dans la répartition des groupements plus ou moins adaptés à chaque variables. Ils ne semblent en effet pas être distribués géographiquement de manière aléatoire, mais apparaissent plutôt comme regroupés en régions sensiblement homogènes. La pertinence d'une variable de second ordre semble donc être liée au contexte climatologique de chaque groupement.

6.2.5 Importance de la qualité de l'analogie de premier ordre

Arrivé à ce stade, nous avons défini un jeu de paramètres qui permet d'obtenir une analogie de second ordre tout à fait intéressante :

- **Variable d'analogie de 2nd ordre** : PW% mesuré à +12h & +24h,
- **Critère d'analogie de 2nd ordre** : $RMSE$,
- **Fenêtre d'analogie de 2nd ordre** : fenêtres de 2.5°x2.5° centrées sur chaque regroupement,
- **Nombre d'analogues retenues** : $N1 = 70$ et $N2 = 30$.

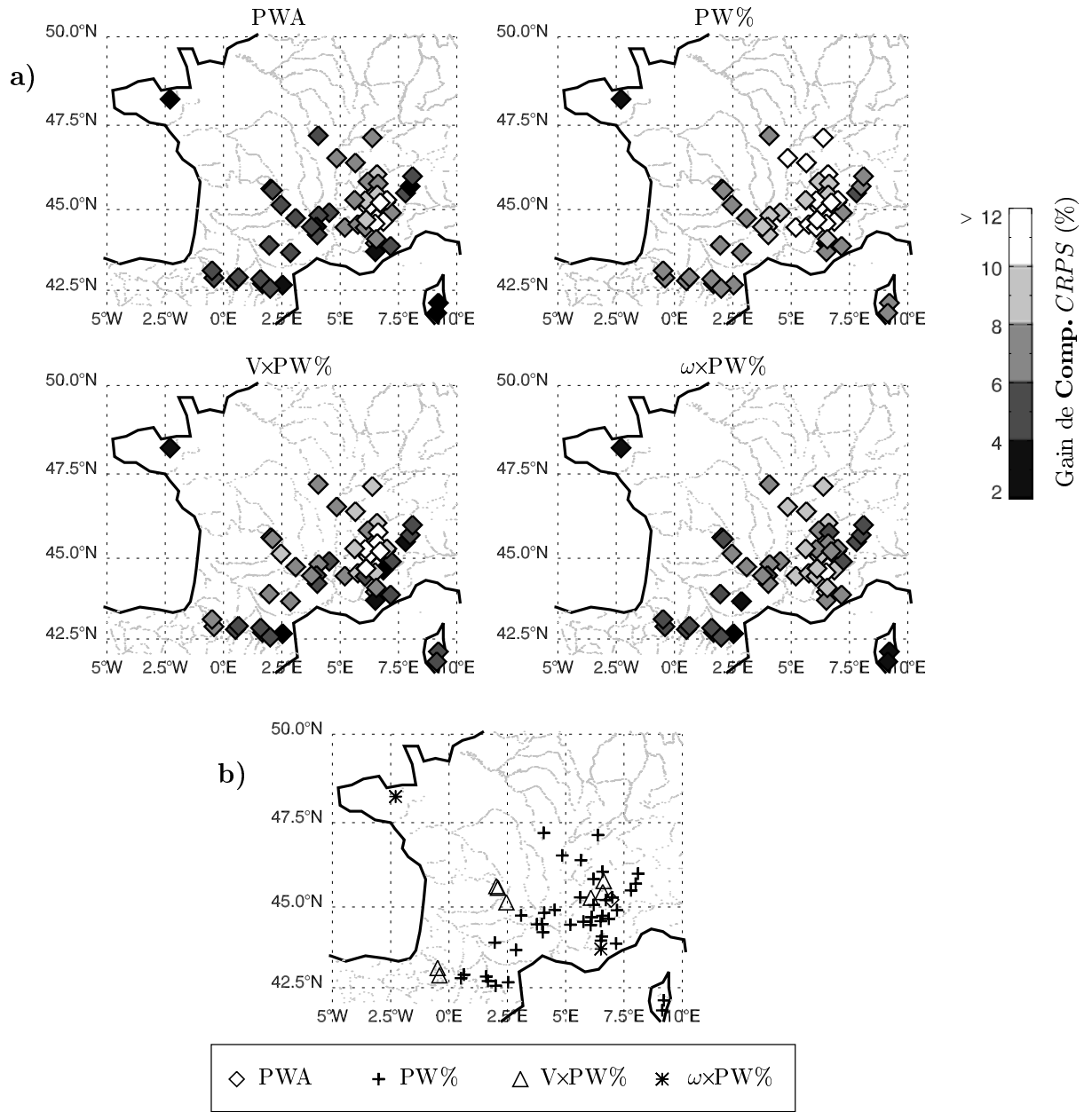


FIG. 6.11: **a)** Gain de performances (en terme de **Comp. CRPS**) obtenu sur chaque groupement en fonction de la variable de second ordre considérée. **b)** Variable de second ordre la plus utile en fonction du groupement considéré.

On peut maintenant se demander à quel point il est utile de raffiner l'analogie de premier ordre pour obtenir des résultats intéressants sur l'analogie de second ordre. Est-ce qu'une amélioration de la reconnaissance des similitudes de circulation synoptique est susceptible de rendre la sélection de second ordre encore plus pertinente ?

Pour tenter de répondre à cette question, nous avons considéré différents algorithmes de sélection de premier ordre, en faisant varier la fenêtre d'analogie et la combinaison de géopotentiels

utilisés. À l'issue de chacune de ces sélections, la même analogie de second ordre est appliquée, ce qui nous permet de suivre l'évolution de l'écart de performance entre les deux sélections, en fonction de la performance intrinsèque de la première sélection (figure 6.12).

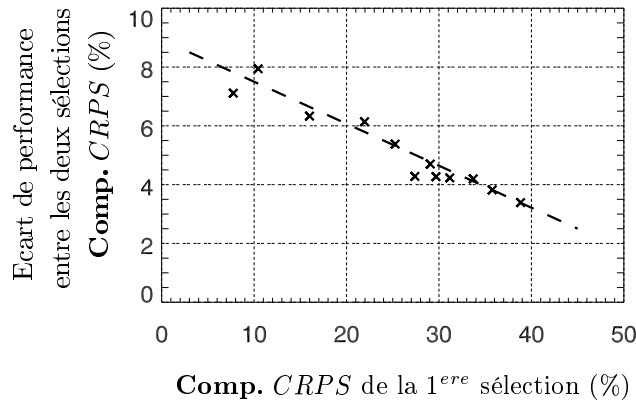


FIG. 6.12: Evolution de l'écart de l'utilité (**Comp. CRPS**) entre les deux niveaux de sélection en fonction de l'utilité de la première sélection (moyennes sur l'ensemble des groupements).

Il apparaît que l'intérêt de l'analogie de second ordre décroît avec l'efficacité de la première sélection. Les limites de performances atteintes par la seconde sélection ne sont donc pas dues à une éventuelle faiblesse dans la première sélection. Ce résultat montre que l'analogie locale sur les variables d'humidité n'a pas forcément besoin d'une analogie synoptique pour apporter de l'information sur les précipitations. Malheureusement, notre graphe ne permet pas de savoir si l'évolution de l'écart suit une asymptote, ou si cet écart devient nul à partir d'un certain seuil d'utilité de la sélection de premier niveau (seuil qu'il n'est peut-être pas possible d'atteindre...). Si ce seuil existait réellement, alors cela signifierait qu'à partir d'une certaine qualité de similitude entre deux situations synoptiques, les valeurs locales de l'humidité, telles que prises en compte ici, n'ont plus la moindre influence sur les précipitations.

En conclusion, il n'est pas utile de raffiner l'analogie de premier ordre à outrance (utilisation de fenêtres spécifiques par types de temps par exemple), en espérant ainsi obtenir un bond quantitatif de performance à l'issue d'une sélection de second ordre.

6.3 Approche dynamique

Dans le chapitre précédent, nous nous sommes demandés s'il était utile de considérer des fenêtres d'analogie différentes en fonction des caractéristiques de la situation cible. Nous avons ainsi défini une approche dynamique qui, même si elle n'apporte que de modestes gains de performance, reste intéressante.

De même, nous pouvons nous demander s'il convient de choisir la variable d'analogie de second ordre en fonction des conditions de la situation cible. En effet, on peut par exemple imaginer que certaines situations, telles les entrées maritimes, présentent des masses d'air dont l'humidité

relative est toujours importante. La différence dans la production de précipitations vient alors plutôt de la présence ou non de mouvements verticaux dans les basses couches. A l'inverse, en ce qui concerne les retours d'Est sur les bassins français des Alpes du Nord, la différence réside dans le caractère plus ou moins humide des masses d'air : la topographie se charge de « *faire précipiter* » ce qui est disponible.

En considérant donc chacun des 6 types de temps définis précédemment (cf. 5.2.3.2), les quatre variables PWA, PW%, V×PW% et $\omega \times \text{PW}\%$ ont été utilisées successivement comme variable de second ordre. L'influence de chaque type de temps sur les groupements diffère évidemment selon la position géographique et le régime climatique de ces derniers. C'est pourquoi il convient de s'intéresser ici de manière individuelle à chaque groupement, ainsi que le fait la figure 6.13, qui présente pour chacun d'eux la variable la plus pertinente en moyenne, en fonction du type de temps considéré.

Il apparaît que le choix de la variable d'analogie de second ordre est bien dépendant de la situation cible. Cette dépendance semble plus ou moins marquée en fonction des groupements. Ainsi, les groupements corses préfèrent toujours PW%, quel que soit le type de temps de la situation cible, alors que les groupements de la Toce reconnaissent alternativement chacune des 4 variables testées comme meilleure variable de second ordre. Par ailleurs, chacune des cartes obtenues présente une certaine cohérence spatiale : les répartitions géographiques des groupements qui préfèrent telle ou telle variable n'apparaissent pas comme aléatoires, mais plutôt organisées de façon régionale. Dans le cas des retour d'Est par exemple, l'ensemble des groupements des Alpes du Nord privilégient PWA, alors que d'une manière générale ils sont plus sensibles à PW% ou V×PW% (cf. figure 6.11b).

Il semble donc que le choix de la variable de second ordre soit guidé par le régime climatique de chaque groupement, c'est à dire par la façon dont ils réagissent à la circulation générale. Ainsi, si l'on s'intéresse au régime de Nord-Ouest, on remarque que les groupements alpins situés *au vent* se comportent de façon identique, mais néanmoins différemment de ceux situés *sous le vent*. Les premiers sont concernés par le flux de PW%, alors que les seconds ne considèrent plus que PW% : au passage des Alpes, le vent dans les basses couches devient fortement perturbé et n'est plus un facteur explicatif.

Mais la différenciation de la variable de second ordre en fonction du type de temps des situations cibles permet-elle pour autant d'obtenir un gain substantiel de performance ?

Le tableau 6.5 présente les résultats détaillés de ces essais pour deux groupements différents : **Gard-Cèze** et **Ht_Agout-Arn-Dourdou**. Ces résultats illustrent les variations de performances que peuvent connaître les différentes variables de second niveau en fonction des types de temps, comme sur le groupement **Gard-Cèze**, où V×PW% n'apporte aucune information dans les situations de retour d'Est (\mathbb{T}_4), alors que cette variable est la plus pertinente pour les situations anticycloniques (\mathbb{T}_6 , avec plus de 12% de gain relatif par rapport à la première sélection).

Ces deux exemples montrent toutefois que PW% demeure globalement la variable de second ordre la plus pertinente. Les gains apportés par le choix de la variable de second niveau en fonction

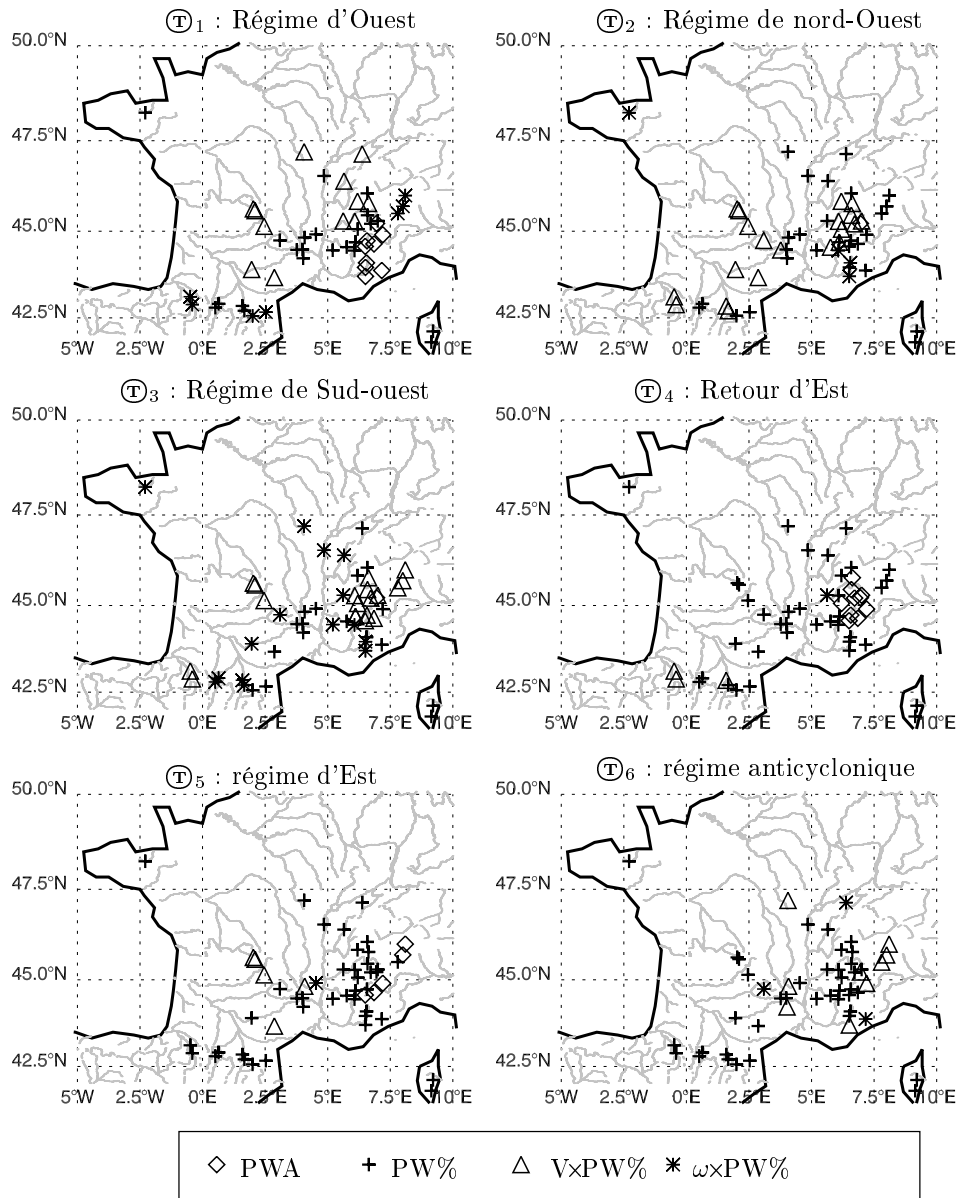


FIG. 6.13: Variable d'analogie de second ordre la plus pertinente pour chacun des groupements, en fonction du type de temps des situations cibles.

GROUPEMENT : Gard-Cèze						
Type	PWA	PW%	<i>Gain sur 1^{ère} sélection</i> V×PW%	<i>Gain sur 1^{ère} sélection</i> ω ×PW%	Meilleure	<i>Gain sur</i> PW%
Ⓓ ₁	6.08	8.53	3.53	4.27	8.53	-
Ⓓ ₂	7.86	13.49	7.50	10.98	13.49	-
Ⓓ ₃	3.37	9.33	6.53	6.31	9.33	-
Ⓓ ₄	6.59	9.79	0.22	3.62	9.79	-
Ⓓ ₅	3.24	6.93	4.40	5.53	6.93	-
Ⓓ ₆	1.96	12.11	12.56	8.74	12.56	0.40
Tous	4.42	9.62	5.45	6.05	9.69	0.05

GROUPEMENT : Ht_Agout-Arn-Dourdou						
Type	PWA	PW%	<i>Gain sur 1^{ère} sélection</i> V×PW%	<i>Gain sur 1^{ère} sélection</i> ω ×PW%	Meilleure	<i>Gain sur</i> PW%
Ⓓ ₁	4.58	6.57	7.03	5.87	7.03	0.43
Ⓓ ₂	3.37	5.48	9.68	2.54	9.68	3.98
Ⓓ ₃	3.05	6.97	5.05	2.77	6.97	-
Ⓓ ₄	8.10	10.88	4.56	3.66	10.88	-
Ⓓ ₅	3.68	8.02	8.59	5.64	8.59	0.53
Ⓓ ₆	2.91	7.61	4.39	3.42	7.61	-
Tous	4.15	7.30	6.27	3.48	8.48	1.10

TAB. 6.5: Gain relatif de performance (**Comp. CRPS**) en fonction du choix de la variable de second niveau et du type de temps de la situation cible. *Gain sur 1^{ère} sélection* : gain relatif par rapport aux performances à l'issue de la 1^{ère} sélection; *Gain sur PW%* : gain relatif par rapport à l'utilisation de PW% comme variable de second niveau.

de la situation cible restent très modestes vis à vis des performances globales. Sur l'ensemble des bassins, on passe ainsi :

- à 42.22% en utilisant uniquement PW%,
- à 42.27% en utilisant la meilleure variable, tous types de temps confondus, pour chaque groupement,
- à 42.46% en utilisant pour chaque type de temps et pour chaque groupement la meilleure variable.

En considérant chaque type de temps individuellement, les gains relatifs peuvent par contre s'élever jusqu'à 3 ou 4%, à l'instar du groupement **Ht_Agout-Arn-Dourdou**, qui voit ses performances en régime de Nord-Ouest passer de 44.47 à 46.24% grâce à l'utilisation de V×PW%. Mais ces gains sont très variables d'un groupement à l'autre.

6.4 Conclusions

Nous avons défini dans ce chapitre une méthodologie qui permet d'enrichir l'analogie grâce à la prise en compte d'une variable atmosphérique supplémentaire. Son application a montré qu'il était intéressant d'ajouter à l'analogie sur la circulation générale, effectuée à une échelle synoptique, une analogie à une échelle plus locale, portant sur une variable (la variable de second ordre) plutôt représentative des processus physiques associés aux hydrométéores.

L'utilisation des variables brutes de l'archive météorologique LTHE a montré que les variables de second ordre les plus informatives sont celles liées à la mesure de la vapeur d'eau atmosphérique (eau précipitable PWA et humidité relative RHU), ainsi que la vitesse verticale du vent (VVE) qui indique les mouvements verticaux de l'atmosphère (convection, soulèvements orographiques...). Ces variables sont donc directement liées aux processus générateurs de précipitations. La construction et l'application de variables plus élaborées, a priori plus représentatives des phénomènes physiques qui nous intéressent, nous a permis d'identifier PW% (le produit de PWA par RHU) comme étant, d'un point de vue global sur l'ensemble des groupements, la variable de second ordre la plus pertinente vis à vis des précipitations.

En étudiant plus finement l'intérêt de ces variables en fonction du groupement considéré et du type de temps de la situation cible, il est apparu que l'utilité de telle ou telle variable dépend du contexte climatique local de chaque situation cible. Ainsi, chaque groupement est susceptible de connaître une variable de second ordre optimale différente pour chaque type de temps identifié, tout en respectant une certaine cohérence régionale. Ces différences sont cependant assez peu sensibles sur les performances globales et ne sont significatives que pour quelques groupements.

L'utilisation de variables de second ordre permet donc un réel gain de performance dans l'explication des précipitations par rapport à une analogie portant uniquement sur la circulation générale. Néanmoins, une grande partie reste encore inexplicée. Evidemment, tout n'est pas explicable (incertitudes de mesures, ...) et la prévision parfaite reste du domaine de l'idéal, mais nous ne pensons pas avoir atteint les limites de prédictibilité.

Outre le problème de la longueur de l'archive qui peut nous restreindre dans le choix des situations candidates, il est fort probable que nous soyons maintenant limités par la résolution spatiale de notre archive. En effet, le caractère local de PW% a été confirmé par l'optimisation de l'algorithme d'analogie : les fenêtres d'analogie optimales se réduisent à seulement 4 points de la grille de l'archive NCEP/NCAR ($2.5^\circ \times 2.5^\circ$). La résolution dont nous disposons ne permet donc pas un échantillonnage spatial suffisant vis à vis des dimensions de la zone pertinente. En outre, on peut se poser des questions quant à la représentativité même de telles variables avec la troncature T62 (≈ 210 km) utilisée par le modèle des réanalyses. Nos données correspondent donc plus ou moins à des valeurs moyennes sur une maille élémentaire, ce qui ne présente pas d'inconvénients pour des variables telles que les hauteurs géopotentielles (HGT), mais peut masquer par exemple la présence d'un jet de basse couche, d'un lac d'air froid dans une vallée alpine, etc.

Les résultats obtenus ici, mettant en lumière le rôle de PW%, sont donc dépendants de l'archive utilisée. Si nous avions disposé d'une meilleure résolution, il est fort possible que les gains de performances aient été plus importants, mais aussi que la hiérarchie entre variables de second ordre ait été modifiée, certaines variables devenant plus pertinentes avec l'augmentation de leur résolution.

L'optimisation des algorithmes a par ailleurs montré que, de même que pour la variable d'analogie de premier ordre, l'heure de mesure de celle de second ordre doit être incluse dans la période de cumul du prédictand. Les variables de second ordre doivent donc elles aussi être prévues par un modèle NWP. Or, ce sont des variables de type B, très influencées par le modèle NWP utilisé,

et dont les prévisions sont moins robustes que celles des variables liées à la circulation générale. Dans le cadre d'une utilisation opérationnelle, la prédictibilité de ces variables à moyen terme peut donc imposer une limitation à la pertinence d'un second niveau d'analogie.

Troisième partie

Etudes diagnostiques et applications



Validation & critique des performances

« *Nam parvolum differt, patiaris adversa an exspectes ;
nisi quod tamen est dolendi modus, non est timendi.
Doleas enim quantum scias accidisse,
timeas quantum possit accidere. »*

« *L'attente du mal ressemble fort au mal lui-même,
à cette différence près que la souffrance a une limite, tandis que la crainte n'en a pas.
On souffre en effet de tout ce qu'on sait être arrivé,
mais on craint tout ce qui peut arriver. »*

Pline le Jeune, Lettres (VIII, 17), traduction de Yves Hucher.

Tout au long de la partie précédente, nous avons cherché quelle était la meilleure façon de sélectionner des situations météorologiques analogues, dans le but d'expliquer les précipitations observées sur des groupements pluviométriques cibles. Ce travail nous a finalement amené à définir un certain nombre d'algorithmes d'identification d'analogies qui se sont révélés optimaux dans leurs contextes.

Cependant, il est toujours possible que les jeux de paramètres conduisant aux meilleures performances ne soient que le fruit du hasard, propre à l'échantillon d'évènements utilisé. La question se pose donc : les résultats obtenus sont-ils reproductibles sur un échantillon indépendant de celui ayant servi au calage des algorithmes ?

C'est pourquoi ce chapitre s'applique à effectuer la validation des résultats de la partie II. Nous en profiterons ensuite pour quantifier les performances de nos adaptations de manière plus fine que ce que nous avons fait jusqu'à présent. Nous utiliserons ainsi des scores qui permettent d'appréhender les performances en fonction de seuils de précipitations ou du choix d'un quantile particulier. Ces considérations nous amèneront finalement, et logiquement, à nous intéresser au biais des prévisions émises et à son éventuelle correction.

7.1 Validation

7.1.1 Méthodologie

Comme annoncé dans l'*introduction à la Partie II*, nous n'avons pas utilisé toutes les données disponibles lors de la phase de calibration des algorithmes. Sur une archive allant de 1953 à 2001, nous n'avons considéré pour l'instant que des situations cibles comprises entre 1953 et 1996 (soit 44 années), nous réservant celles allant de 1997 à 2001 (soit 5 années) pour la phase de validation.

Dans les deux cas néanmoins, afin que les résultats demeurent comparables, la période autorisée pour la sélection des situations candidates est la même : de 1953 à 1996. Etant donné que l'on s'interdit d'accepter une situation candidate appartenant à la même année que la situation cible, cela correspond à une archive effective de 43 années pour la phase de calibration, et de 44 pour celle de validation.

Nous sommes ainsi en mesure de *valider* les algorithmes définis lors de la phase de calibration, en testant leur pertinence sur un échantillon de situations cibles indépendant de celui déjà utilisé. Cette validation ne porte toutefois pas sur les bassins italiens, les données de précipitations n'ayant pas pu être mises à jour au delà de 1998.

Les performances sont ici encore exprimées en terme de compétence du *CRPS* (**Comp. CRPS**) par rapport aux performances de la distribution climatologique sur le même échantillon. L'archive dans laquelle sont piochées les situations candidates étant identique pour la calibration et la validation, la distribution climatologique considérée est elle aussi la même, à savoir la distribution des pluies journalières entre 1953 et 1996.

7.1.2 Algorithmes retenus

A l'issue des différentes approches testées dans la reconnaissance de situations analogues, nous proposons de retenir 3 algorithmes différents : M1, M2 et M3, tels que définis par le tableau 7.1. Ceux-ci correspondent à des niveaux de raffinement croissants de l'analogie mais aussi à une augmentation des temps de calculs et des volumes de données nécessaires à leur mise en œuvre. Par ailleurs, afin de les situer par rapport à ce qui existe déjà, nous considérons aussi l'algorithme de référence M0 issu des travaux de S. Guilbaud (1997), lequel est utilisé de façon opérationnelle et nous a servi de point de départ.

La principale différence entre M1 et M0 réside dans l'utilisation de champs de géopotentiels mesurés à +12h au lieu de +00h, ce qui permet, pour des performances équivalentes (voire légèrement supérieures), de réduire le nombre de champs utilisés de 4 à 2. Ceci est possible grâce à l'utilisation des réanalyses qui rendent disponibles ces mesures à +12h. M1 est donc un algorithme plus simple que M0.

M2 est basé sur les mêmes paramètres que M1, mais diffère par la fenêtre d'analogie utilisée. Les groupements sont effectivement rassemblés en 7 regroupements régionaux $\textcircled{\mathbf{R}}_1$ à $\textcircled{\mathbf{R}}_7$, comme défini dans le chapitre V (cf. 5.1.4.2), et au lieu d'utiliser une fenêtre d'analogie commune, on

		ALGORITHMES DE SÉLECTION			
		M0	M1	M2	M3
<i>Première sélection</i>	Variable	HGT	HGT	HGT	HGT
	Niveau et heure de mesure	z1000+00 z1000+24 z700+00 z700+24	z1000+12 z500+24	z1000+12 z500+24	z1000+12 z500+24
	Fenêtre	lon0 = -10°E lat0 = 37.5°N dlon = 25° dlat = 12.5°	lon0 = -7.5°E lat0 = 37.5°N dlon = 25° dlat = 10°	fenêtres propres à chaque regroupement (\mathbb{R}_1 à \mathbb{R}_7)	fenêtres propres à chaque regroupement (\mathbb{R}_1 à \mathbb{R}_7)
	Critère	S1	S1	S1	S1
	N1	50	50	50	70
<i>Seconde sélection</i>	Variable	-	-	-	PW%
	Niveau et heure de mesure	-	-	-	z850+12 z850+24
	Fenêtre	-	-	-	fenêtres de 2.5x2.5° centrées sur chaque regroupement (\mathbb{R}_1 à \mathbb{R}_7)
	Critère	-	-	-	RMSE
	N2	-	-	-	30

TAB. 7.1: Description des algorithmes de sélection de situations analogues retenus pour la phase de validation.

considère une fenêtre différente pour chacun des regroupements. Ces fenêtres sont celles qui ont été optimisées à l'occasion de la définition des regroupements (cf. tableau 5.2). On ne s'attend donc pas à ce que M2 produise globalement des performances fortement supérieures à M1, mais simplement à ce que celles-ci soient plus homogènes d'un groupement à l'autre.

L'algorithme M3 correspond à l'adjonction d'une analogie de second ordre à M2. La variable de second ordre utilisée est PW% (produit de l'eau précipitable par l'humidité relative à 850 hPa), considérée sur des fenêtres très locales de seulement 4 points de grille (2.5x2.5°), centrées sur chacun des regroupements. Cet algorithme est le plus performant, mais c'est aussi le plus gourmand en données, nécessitant la connaissance de trois variables (HGT, PWA et RHU) à deux horaires de mesure différents (+12h et +24h). Lors d'une utilisation en prévision opérationnelle, cet algorithme sera donc plus sensible aux éventuelles erreurs de prévision du modèle NWP qui doit servir à décrire les situations cibles, d'autant plus que les prévisions des variables mesurant l'humidité sont moins robustes que celles concernant la circulation atmosphérique générale.

7.1.3 Comparaison des groupements

Avant de valider les algorithmes retenus, il peut être utile de s'intéresser aux performances obtenues par M0 sur la période de validation pour chacun des groupements. Cet algorithme

ayant déjà été validé, il sera ainsi possible de vérifier groupement par groupement si les données pluviométriques sont restées homogènes entre la période de calibration et celle de validation.

La figure 7.1 présente ainsi pour chaque groupement les écarts relatifs de performance (**Comp. CRPS**) mesurés entre les périodes de calibration et de validation. Elle montre que globalement, les performances sont plutôt plus élevées en validation qu'en calibration, ce qui est un résultat peu courant dans les problèmes d'optimisation.

Pour expliquer ce gain de performance, on peut essayer d'avancer plusieurs hypothèses :

- Avec le protocole adopté, la validation dispose de 44 années de situations cibles alors que la calibration n'en a que 43. Disposant d'un panel un peu plus important, il est donc plus facile pour la validation de trouver des situations présentant de fortes analogies.

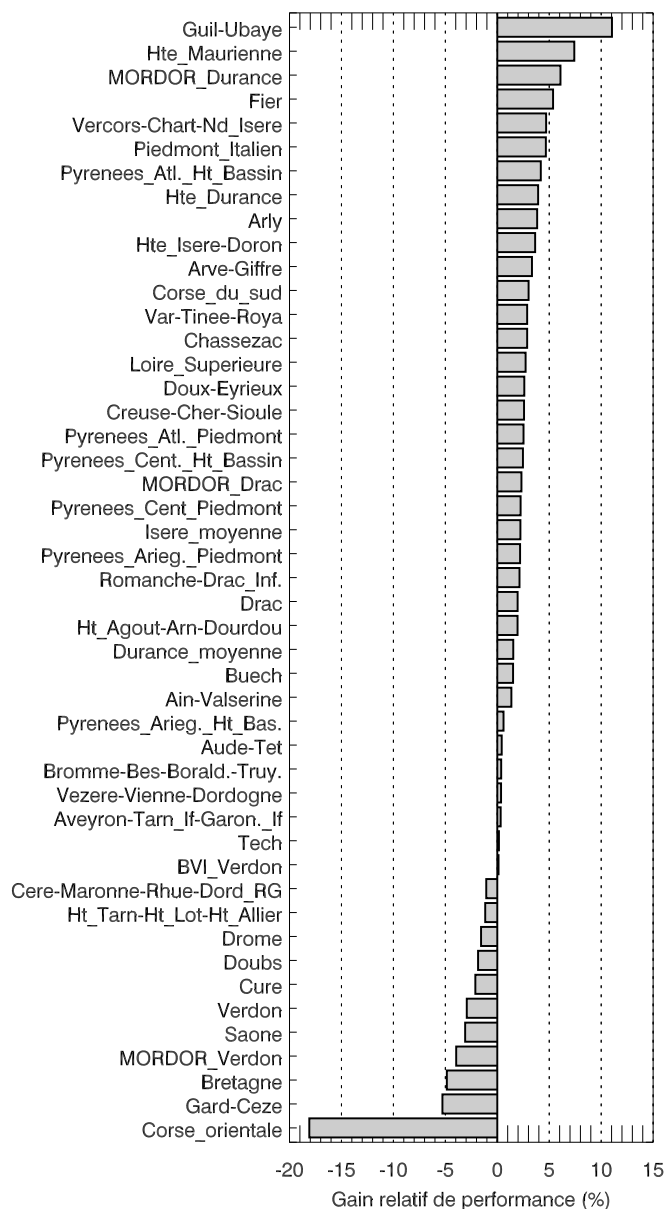


FIG. 7.1: Ecarts relatifs de performance (**Comp. CRPS**) entre période de calibration et période de validation pour l'algorithme M0.

- Des problèmes d'hétérogénéité à l'intérieur de l'archive météorologique, discutés dans le chapitre suivant, peuvent aussi être à l'origine de ce résultat. Il est ainsi possible que ces éventuelles hétérogénéités soient moins sensibles vis à vis de la période de validation que de celle de calibration. Cela aurait pour effet de restreindre artificiellement le panel de situations candidates effectivement disponibles pour les situations cibles de la période de calibration. Ce serait donc les performances de la phase de calibration qui seraient sous-estimées.
- On peut aussi imaginer que la période de validation, relativement courte, présente des situations relativement « *faciles* » à prévoir. Notre mesure des performances est effectivement exprimée en terme de compétence relative par rapport à une prévision basée sur la distribution climatologique. Si la climatologie empirique de la période de validation s'écarte notablement de la distribution de référence, une prévision même moyenne montrera une forte compétence.

Il n'est pas possible de quantifier ici la part respective de chacune de ces explications, mais nous reviendrons plus longuement dans le chapitre suivant sur les deux premiers points. Nous y verrons que la seconde hypothèse est probablement celle à privilégier.

La figure 7.1 montre par ailleurs que le groupement **Corse_orientale** présente un comportement particulier : ses performances chutent lorsque l'on considère la période de validation. M0 ayant déjà été validé lors de travaux précédents (Guilbaud 1997), ceci n'est sûrement pas dû à une sur-optimisation de l'algorithme, mais plus probablement à une rupture dans les données pluviométriques de ce groupement.

Nous avons donc comparé, pour quatre groupements, les distributions empiriques des précipitations mesurées sur les périodes 1953-96 et 1997-2001 (figure 7.2). Il apparaît que les groupements **Gard_Ceze** et **Corse_orientale**, qui affichent les plus importantes pertes de performances en validation, présentent des distributions fort suspectes. Non seulement **Gard_Ceze** connaît une fréquence des valeurs nulles exagérément forte sur la période de validation, mais surtout, ces deux groupements ont enregistré une dizaine d'évènements décennaux ($R \geq 1$) au cours de ces 5 années.

Ceci ne peut s'expliquer que par un improbable concours de circonstances dont nous aurions gardé la mémoire, ou plus sûrement, par une modification des postes pluviométriques entrant dans le calcul des lames d'eau moyennes reçues sur ces groupements. Les données ne sont alors plus comparables entre elles. Si le prédictand n'est pas le même pour les situations cibles et les situations candidates, les prévisions sont forcément peu efficaces.

Les groupements de l'archive EDF ont donc connu des modifications, ce qui nous a été confirmé par la DTG. Malheureusement, le document récapitulant l'ensemble de ces changements ne nous est pas parvenu à l'heure de l'écriture de ces lignes. Néanmoins, dans le cadre de l'utilisation que nous en faisons, ces modifications ne sont fortement sensibles, d'un point de vue global sur toute la période de validation, que pour le groupement **Corse_orientale**.

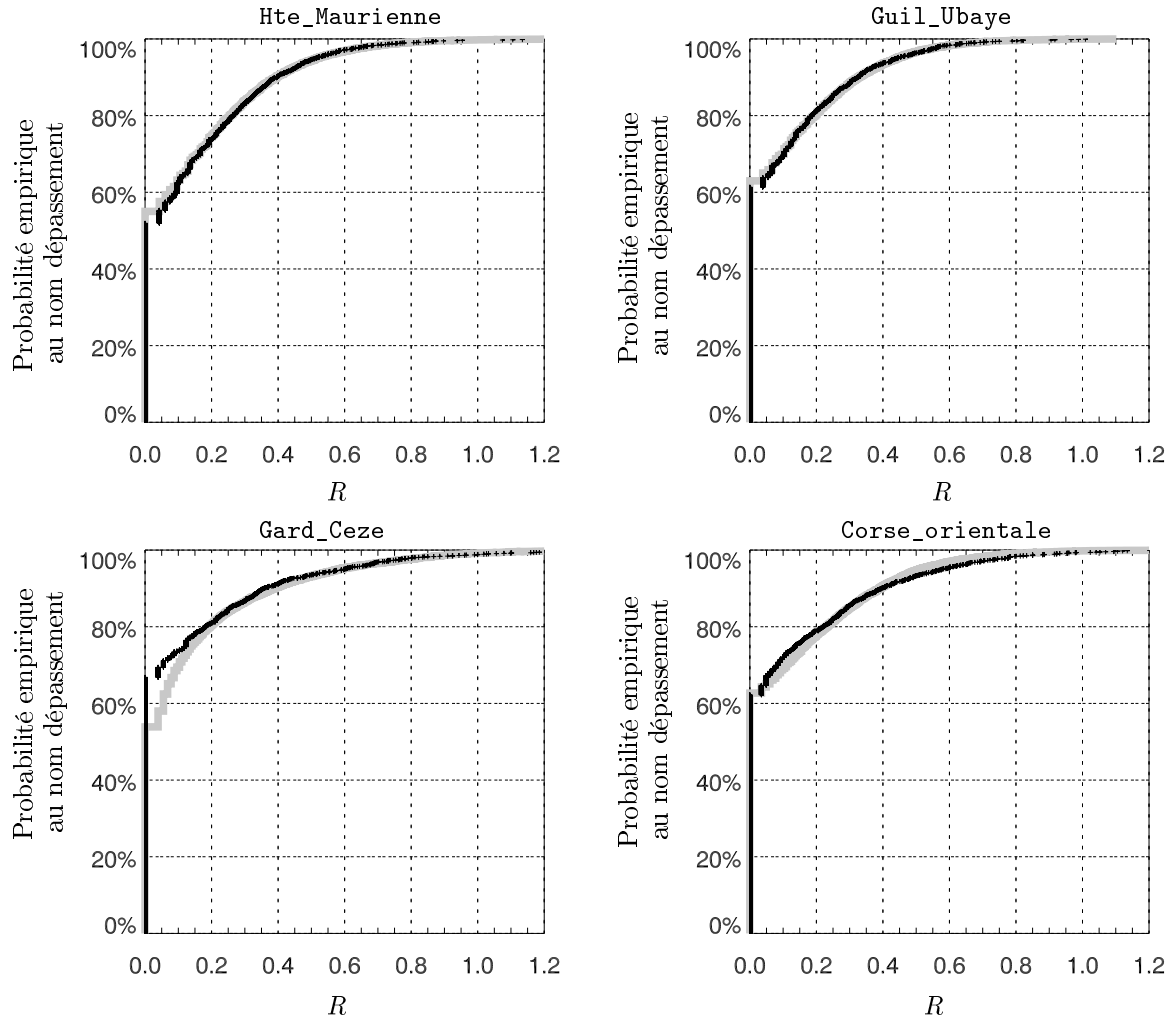


FIG. 7.2: Comparaison des distributions empiriques de précipitation mesurées sur les périodes 1953-96 (courbe grisée) et 1997-2001 (croix noires), pour quatre groupements de l'archive EDF.

7.1.4 Validation des algorithmes

Les performances obtenues pour chacun des quatre algorithmes, selon que l'on considère la période de calibration ou celle de validation, sont présentées figure 7.3. Ces résultats concernent l'ensemble des groupements, *Corse_orientale* y compris.

On constate que la hiérarchie entre les différents algorithmes est parfaitement respectée sur la période de validation : des performances croissantes sont associées dans l'ordre à M0, M1, M2 et M3. L'optimisation effectuée est donc reproductible d'un échantillon à l'autre : elle n'est pas due au hasard et reflète une réalité physique. On peut donc considérer que l'étape de validation est passée avec succès.

De même que pour M0, la validation présente systématiquement de meilleures performances que la calibration pour tous les algorithmes (M1, M2 et M3). Ceci est cependant beaucoup moins sensible pour M2 que pour les autres algorithmes. En effet, les gains qu'apporte M2 sur M1 se concentrent sur quelques groupements excentrés. Or, comme nous venons de le voir, les

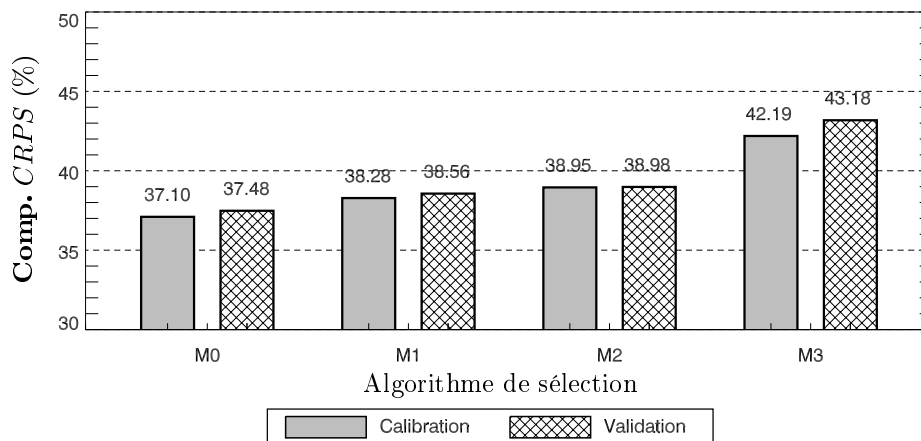


FIG. 7.3: Comparaison des performances moyennes obtenues sur les périodes de calibration et de validation pour chacun des quatre algorithmes retenus.

groupements **Corse_orientale** et **Bretagne**, qui font partie de ceux concernés par M2, semblent ne plus être homogènes par rapport à la période de calibration. L'intérêt de M2 ne peut donc pas apparaître pleinement à la validation.

Au contraire, M3 est l'algorithme qui profite le plus du gain apporté par la validation. L'information locale est donc encore plus pertinente sur la période de validation. Cela pourrait être dû à une meilleure représentation dans les années récentes des variables d'humidité (PWA et RHU) par les réanalyses.

7.2 Etude des performances

Les algorithmes d'analogie ont donc été validés, notamment en ce qui concerne M3, lequel permet un gain de performance significatif par rapport aux autres algorithmes. Néanmoins, cette notion de performance, mesurée via la compétence du *CRPS*, reste relativement floue. Ce score est en effet un intégrateur des qualités et des défauts des prévisions produites. Pour un utilisateur potentiel, il est important de disposer d'une description plus fine des performances réalisées : la prévision est-elle aussi efficace pour prévoir les fortes précipitations que pour prévoir l'occurrence de pluie ? Présente-t-elle plutôt une tendance au défaut d'alerte ou à la fausse alerte ?

Nous allons donc maintenant nous intéresser aux différents aspects de la performance des algorithmes retenus grâce à l'application de quelques uns des scores décrits dans le chapitre 3. Nous nous plaçons ici sur la période de validation et nous écartons le groupement **Corse_orientale** du panel des groupements considérés, à cause des changements intervenus dans sa composition.

Les scores que nous allons utiliser sont ceux basés sur l'utilisation de tableaux de contingence, tels que décrits dans la section 3.2.1.1. Rappelons que, quand on dispose d'une série de prévisions déterministes de l'occurrence ou non d'un événement, un tableau de contingence correspond au recensement du nombre de prévisions faites de l'évènement en fonction de l'occurrence réelle de celui-ci (tableau 7.2).

		Observations	
		Oui	Non
Prévisions	Oui	a	b
	Non	c	d

TAB. 7.2: Exemple de tableau de contingence.

A partir de ce tableau, nous définissons les scores suivants :

- **Probabilité de Détection, POD** : probabilité conditionnelle de faire une bonne prévision d'alerte, sachant qu'une alerte est observée.

$$POD = \frac{a}{a + c}$$

- **Taux de Fausses Alertes, FAR** : probabilité de faire une mauvaise prévision sachant qu'une alerte a été prévue (FAR est orienté négativement).

$$FAR = \frac{b}{a + b}$$

- **Spécificité, SPE** : probabilité conditionnelle d'émettre un 'NON' correct lorsque qu'aucune alerte n'est effectivement observée.

$$SPE = \frac{d}{b + d}$$

On définit par ailleurs le Taux de Fausse Détection comme étant égal à $1 - SPE$.

- **True Skill Score, TSS** : différence d'Indice de Réussite entre la prévision testée et une prévision faite au hasard, rapportée à l'Indice de Réussite obtenu par une prévision faite elle aussi au hasard mais respectant la climatologie observée de l'événement sur l'échantillon considéré.

$$TSS = POD - (1 - SPE)$$

Pour de plus amples détails, le lecteur est invité à se reporter à la section 3.2.1.1 du chapitre 3.

Nous considérons ici qu'un événement est un dépassement d'un seuil de précipitation X_s . Par ailleurs, à partir des fonctions de distributions prévues, nous sommes en mesure d'émettre une prévision déterministe en considérant la valeur associée à une fréquence au non dépassement donnée. Cette valeur est appelée quantile $x\%$ et notée $Q_{x\%}$, pour une fréquence au non dépassement de $x\%$ (figure 7.4). Ainsi, à chaque couple $(X_s ; Q_{x\%})$ est associé un tableau de contingence différent, et donc des scores différents.

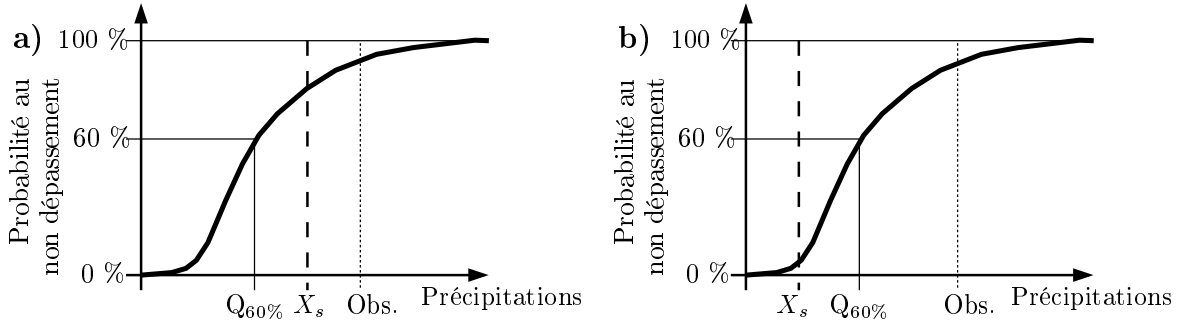


FIG. 7.4: Exemple de prévision probabiliste pour laquelle on s'intéresse au quantile $Q_{60\%}$. **a)** le seuil X_s est situé entre $Q_{60\%}$ et l'observation : cas c (observation = Oui et prévision = Non) ; **b)** le seuil X_s est inférieur à $Q_{60\%}$ et à l'observation : cas a (observation = Oui et prévision = Oui).

7.2.1 Pouvoir discriminant en fonction d'un seuil

Les diagrammes ROC (cf. 3.3.2) permettent de faire apparaître, pour un seuil de précipitations donné, le pouvoir globalement discriminant d'une prévision probabiliste vis à vis des événements qui dépassent ou non ce seuil. Ils se construisent en traçant, pour l'ensemble des valeurs de $Q_x\%$ possibles, la probabilité de détection (POD) en fonction du taux de fausse détection ($1 - SPE$). On obtient ainsi une courbe qui joint :

- le point origine (0;0) : pour $Q_{0\%}$, on ne prévoit jamais le moindre événement ; on ne détecte donc rien ($POD=0$) mais on ne fait pas de fausse détection non plus ($1-SPE=0$),
- au point (1;1) : pour $Q_{100\%}$, on prévoit systématiquement un événement ; on détecte donc tout ($POD=1$) mais s'il ne se produit rien, on fait à coup sûr une fausse détection ($1-SPE=1$).

Dans un tel diagramme (dont un croquis schématique est donné par la figure 3.9, chapitre 3), une prévision basée sur la distribution climatologique de l'échantillon suit la première bissectrice. Une prévision parfaite (prévision déterministe de la valeur observée) passe par sa part par le point (0;1) : probabilité de détection de 100% associé à un taux de fausse détection nul. Ainsi, plus la courbe ROC d'une prévision s'écarte de la première bissectrice pour se rapprocher du point (0;1), meilleure est cette prévision.

La mesure de l'aire sous les courbes ROC fournit un indicateur de la performance, qui va donc de 0.5 pour une prévision qui n'apporte rien de plus que la climatologie à 1.0 pour une prévision parfaite.

La figure 7.5a présente les courbes ROC obtenues sur la période de validation par les 4 algorithmes retenus pour un seuil pluie/non-pluie. On constate que ces courbes s'écartent notablement de la première bissectrice : les prévisions possèdent un réel pouvoir discriminant. Les aires situées sous ces courbes dépassent ainsi 0.880 . Elles présentent par ailleurs un comportement équilibré entre capacité de détection et risque de fausse détection. Cependant, de même que pour la compétence du $CRPS$, les comportements de M0, M1 et M2 se distinguent peu les uns des autres. Seul M3 se détache du lot en affichant un pouvoir discriminant supérieur (aire de 0.898).

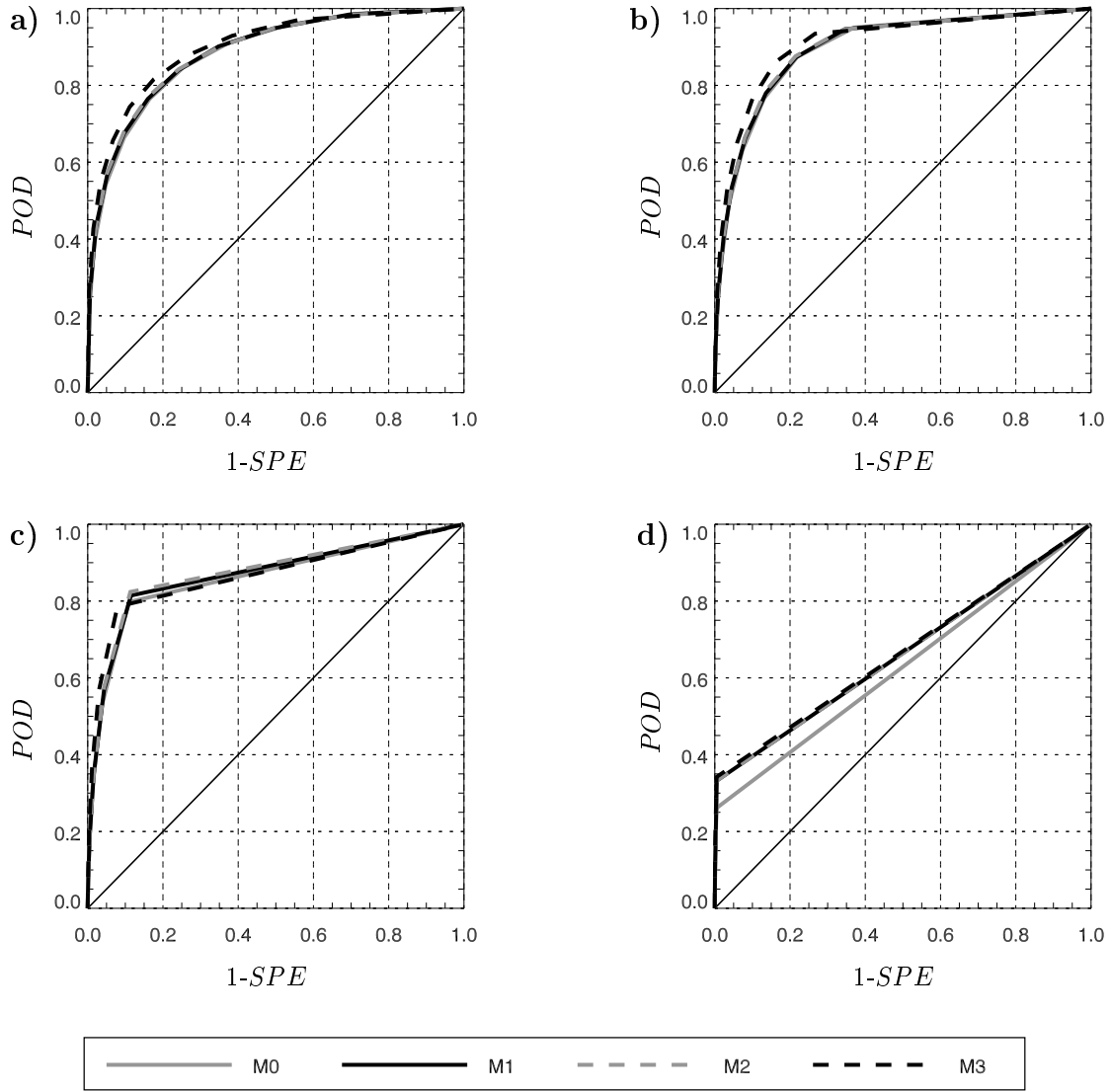


FIG. 7.5: Courbes ROC sur la période de validation pour les 4 algorithmes retenus. **a)** : seuil pluie/non-pluie; **b)** : seuil $R = 0.25$ ($R_{brute} = R_{10}/16$); **c)** : seuil $R = 0.50$ ($R_{brute} = R_{10}/4$); **d)** : seuil $R = 0.75$ ($R_{brute} = R_{10}/1.78$).

Si l'on considère un seuil de pluie légèrement supérieur au seuil pluie/non-pluie (figure 7.5b, seuil $R_{brute} = R_{10}/16$), les prévisions se révèlent encore plus discriminantes, avec des aires de l'ordre de 0.895 pour M0, M1 et M2, et montant jusqu'à 0.910 pour M3. On se situe ici vers le seuil de prévision optimal. Les courbes présentent par contre une très légère dissymétrie, qui indique que les prévisions auraient tendance à privilégier la détection, au risque d'augmenter le taux de fausses détections.

En s'intéressant maintenant à un seuil de précipitations relativement fortes (figure 7.5c, seuil $R_{brute} = R_{10}/4$), on constate un changement dans le comportement des courbes ROC. Le pouvoir discriminant des prévisions reste important (aires de l'ordre de 0.865), mais :

- d'une part la supériorité de M3 par rapport aux autres algorithmes est moins marquée,

- et d'autre part, les courbes montrent clairement que le faible taux de fausses détections prime devant la capacité de détection. Ces prévisions « *ratent* » des événements, mais crient rarement au loup pour rien.

Cette tendance s'accroît ensuite nettement au fur à mesure que l'on considère des seuils de précipitations importants. Ainsi, pour les événements relativement rares (figure 7.5d, seuil $R_{brute} = R_{10}/1.78$), l'aire sous les courbes chute aux alentours de 0.663 et la probabilité de détection devient très faible devant le taux de fausses détections. Il est intéressant de noter que, pour ces fortes précipitations, M3 est revenu au même niveau de performance que M2 et M1, alors que M0 est nettement distancé (aire de 0.628). La prise en compte de champs de géopotentiels mesurés à +12h permet donc d'améliorer la détection des forts événements, alors que celle des variables locales d'humidité (analogie en deux temps) favorisait plutôt la prévision des faibles seuils.

En conclusion de l'étude des courbes ROC, on peut dire que les algorithmes de prévision par adaptation possèdent une importante capacité de discrimination, avec un optimum centré sur les faibles précipitations plus que sur la détection pluie/non-pluie. Néanmoins, pour les fortes précipitations ($R_{brute} > R_{10}/4$), leurs performances chutent assez rapidement : elles gardent de très faibles taux de fausses détections mais voient s'affaiblir leur capacité de détection.

7.2.2 Choix d'un quantile de référence

Dans le cadre d'une utilisation pratique, il peut être pratique d'exprimer une prévision probabiliste sous une forme déterministe. On dispose ainsi d'une unique valeur plus facile à appréhender. Pour ce faire, quel quantile est-il le plus judicieux de choisir au sein de toute la distribution ?

Nous allons donc nous intéresser à l'évolution des scores issus des tableaux de contingence en fonction du quantile $Q_x\%$ considéré. La figure 7.6a présente ainsi les scores *POD*, *FAR* et *TSS* obtenus sur la période de validation pour les quantiles de M3 allant de $Q_{10}\%$ à $Q_{90}\%$, par tranche de 10%. A titre d'exemple de lecture de ces graphes, si l'on choisit un seuil de précipitation $R = 0.6$, on constate que le quantile $Q_{90}\%$ de M3 affiche un *POD* de 0.62, c'est à dire que ce seul quantile $Q_{90}\%$ a permis d'anticiper 62 % des événements qui ont engendré des précipitations telles que $R \geq 0.6$.

Il apparaît que les différents quantiles présentent une forte dispersion au niveau de *POD*, allant d'un maximum de 0.27 pour $Q_{10}\%$ à 0.97 pour $Q_{90}\%$. Fort logiquement, les faibles quantiles détectent peu d'événements, tandis que les plus élevés n'en laissent quasiment pas passer. Nous retrouvons ici le manque de finesse des prévisions par analogie. Toutefois, cette dispersion est moins prononcée en ce qui concerne *FAR*. Le fort taux de détection des quantiles élevés n'est donc pas simplement dû au fait que ceux-ci maximisent la prévision, mais est plutôt lié à un réel pouvoir de prévision.

Le score *TSS*, qui intègre aussi bien la capacité de détection que l'aptitude à éviter les fausses détections, permet de quantifier la performance globale de chaque quantile. On constate ainsi que selon le seuil de précipitation auquel on s'intéresse, il convient de ne pas s'attacher au même

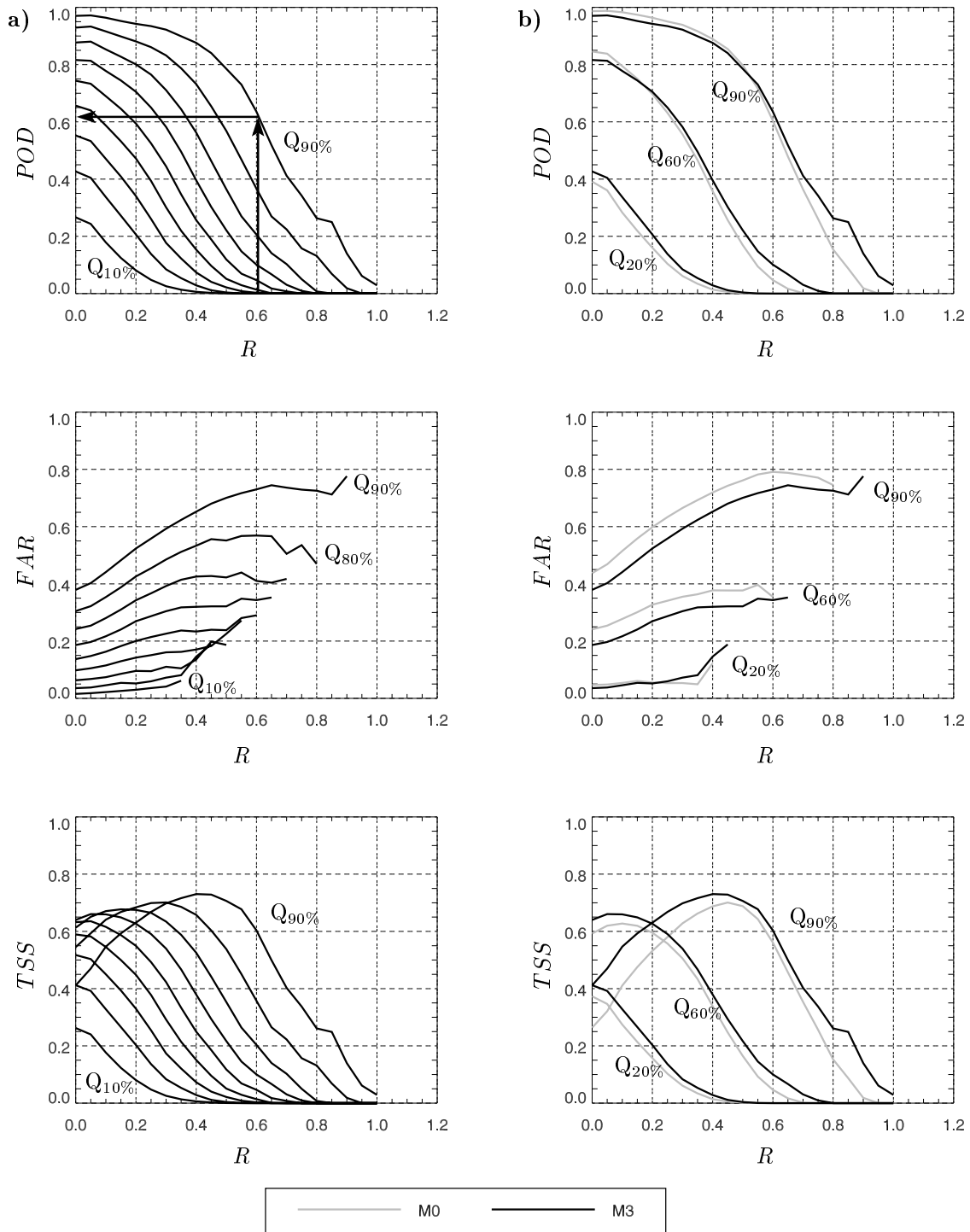


FIG. 7.6: Evolution des scores POD , FAR et TSS sur la période de validation en fonction de R . **a)** : quantiles $Q_{10\%}$ à $Q_{90\%}$ par tranche de 10% pour M3; **b)** : comparaison entre M0 et M3 pour les quantiles $Q_{20\%}$, $Q_{60\%}$ et $Q_{90\%}$.

quantile. Ainsi, pour prévoir l'occurrence de précipitation (seuil pluie/non-pluie), le quantile le plus informatif est $Q_{60\%}$, mais dès lors que l'on s'intéresse à des quantités de pluies telles que $R \geq 0.3$ ($\approx R_{brute} > R_{10}/10$), il est préférable de considérer $Q_{90\%}$.

Les quantiles inférieurs à $Q_{60\%}$ contiennent peu d'informations. Le principal intérêt des faibles quantiles, à l'image de $Q_{20\%}$, est leur très faible taux de fausse alerte, presque constant pour tous les seuils de précipitations. Ils peuvent donc être considérés comme une limite inférieure de la prévision.

Cependant, quel que soit le quantile ou le score considéré, les performances s'effondrent pour les seuils de précipitations les plus élevés. Du fait des limitations de l'archive, qui ne contient nécessairement que peu d'événements extrêmes, il n'est pas possible d'extraire 30 observations de pluies décennales par exemple. Les prévisions connaissent donc une limite d'un point de vue quantitatif, même si elles demeurent informatives d'un point de vue qualitatif.

La figure 7.6b permet de comparer les performances de M0 et de M3 sur les quantiles $Q_{20\%}$, $Q_{60\%}$ et $Q_{90\%}$. *TSS* montre que pour tous les seuils et tous les quantiles, M3 est effectivement plus performant que M0. Néanmoins l'origine de ce gain de performance n'est pas le même selon le seuil de précipitation considéré :

- pour les seuils faibles, M3 présente des performances légèrement inférieures à celles de M0 en ce qui concerne la probabilité de détection *POD*. Le gain apporté par M3 réside ici plutôt dans une diminution sensible du taux de fausse alerte *FAR*.
- pour les seuils plus importants, cette diminution de *FAR* est moins importante, mais elle se conjugue à un gain de *POD*.

7.2.3 Exemple de prévisions

Les scores que nous venons d'étudier nous apportent de précieuses indications sur les points forts et les faiblesses des prévisions issues de nos algorithmes d'analogie. Toutefois, cette vision reste celle d'un comportement moyen. Chaque prévision est un cas particulier et il est donc primordial de s'intéresser à de vrais exemples de prévisions.

Les deux exemples que nous montrons ici prennent place dans la période de validation, c'est à dire que ces événements n'ont pas été utilisés dans la phase d'optimisation des algorithmes.

Exemple 1 : groupement *Bromme_Bes_Borald._Truy*.

Nous présentons ainsi dans la figure 7.7 les prévisions émises du 4 décembre 1997 au 16 janvier 1998 sur le groupement *Bromme_Bes_Borald._Truy*. par les algorithmes M0 et M3. Ces prévisions, exprimées sous la formes des trois quantiles $Q_{20\%}$, $Q_{60\%}$ et $Q_{90\%}$, sont comparées avec les observations relevées sur la même période. Nous nous plaçons néanmoins toujours ici dans un contexte de prévision parfaite.

Cet exemple montre qu'il y a une bonne co-fluctuation entre les prévisions et les observations, aussi bien pour M3 que pour M0. L'algorithme M3 présente toutefois l'avantage d'une dispersion moindre de ses prévisions :

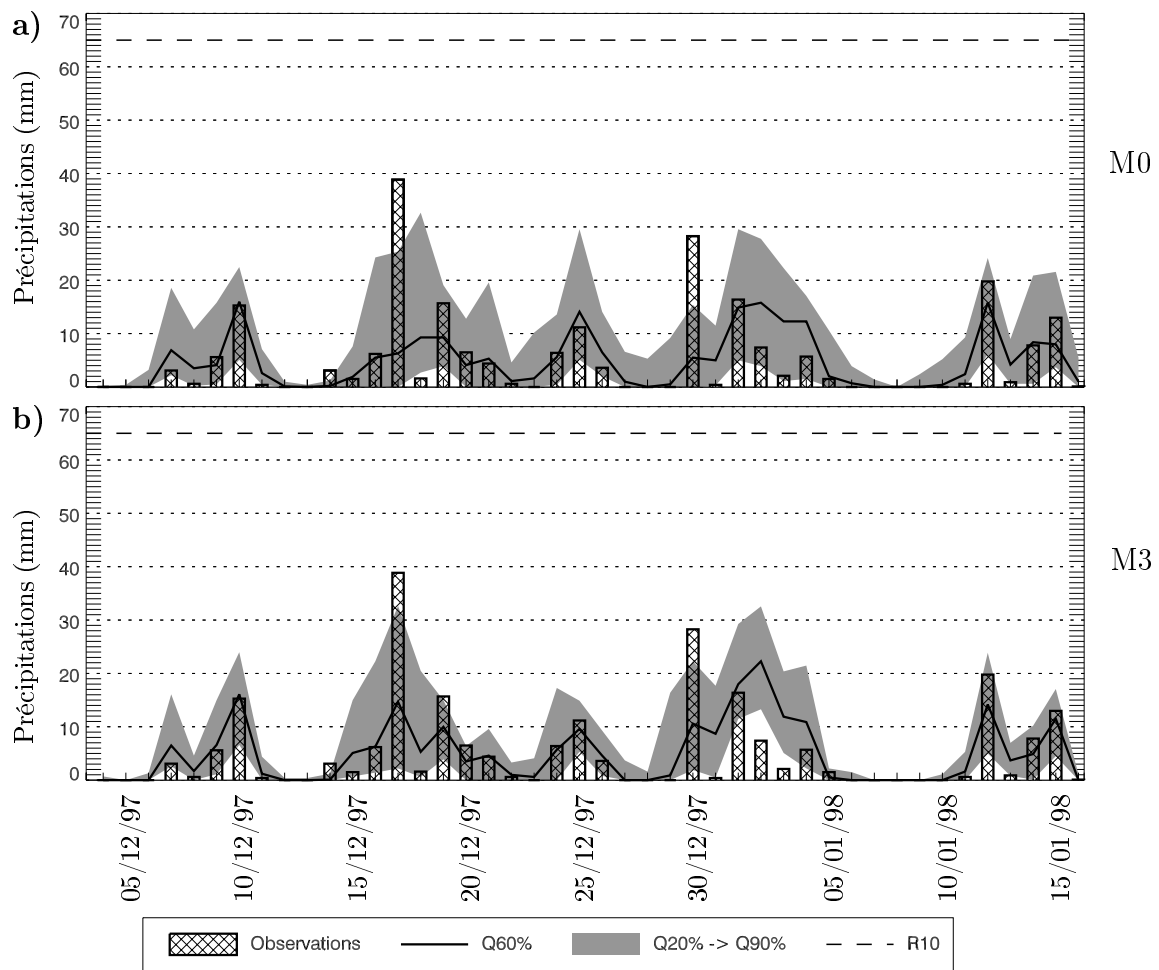


FIG. 7.7: Comparaison sur le groupement Bromme_Bes_Borald._Truy., du 4/12/97 au 16/01/98, entre les précipitations observées et prévues par les algorithmes M0 (a)) et M3 (b)), en prévision parfaite.

- aussi bien dans le cas de la reconnaissance des journées non pluvieuses, qui se fait de manière plus assurée, comme le montre par exemple la période du 6 au 10 janvier 1998,
- que dans celui de la prévision de précipitations moyennes, comme l'illustre l'épisode du 24 au 26 décembre 1997.

Dans le cas des événements plus importants, M3 améliore aussi les performances de M0. Ainsi, sur l'épisode du 17 décembre 1997, qui a engendré une précipitation observée de 39 mm (avec $R_{10} = 66$ mm), le $Q_{60\%}$ voit sa valeur doubler d'un algorithme à l'autre. Par ailleurs, M3 situe bien le maximum de précipitation sur le 17 décembre alors que M0 le décale sur le lendemain. Il en est de même lors de l'événement du 30 décembre (28 mm observé), pour lequel $Q_{90\%}$ passe de 15 à 22 mm, tandis que le $Q_{60\%}$ double de 5 à 11 mm. Cependant, les prévisions restent systématiquement en deçà des observations pour ces événements relativement forts. On retrouve ici la chute de performance que montraient les scores précédents, due nous semble-t-il aux limitations des archives utilisées.

Bien que M3 présente globalement un avantage certain sur M0, on constate quelques ratés de l'analogie en deux temps. C'est notamment le cas lors des 2 et 3 janvier 1998, qui ont vu des précipitations moyennes à faibles, de respectivement 8 et 3 mm :

- M0 fournit pour ces deux journées une fourchette de prévision très large, mais qui inclut quand même les observations.
- M3 réduit sensiblement la fourchette de prévision, mais malheureusement dans le mauvais sens. Les observations sont fortement sur-estimées, même par le quantile $Q_{20\%}$.

Ces fausses alertes illustrent la sensibilité accrue de l'analogie en deux temps à la description météorologique des situations cibles. Il est en effet fort probable que, soit du fait d'une mauvaise localisation (due à la résolution insuffisante des réanalyses), ou soit à cause d'imprécisions dans les données, l'information hygrométrique ait été surestimée. La seconde sélection a ainsi privilégié les situations candidates présentant de fortes valeurs de PW%, et donc de forts cumuls de précipitations, alors que des situations candidates qui correspondaient plus aux observations et qui étaient présentes à l'issue de la première sélection ont été écartées.

Exemple 2 : groupement Arly

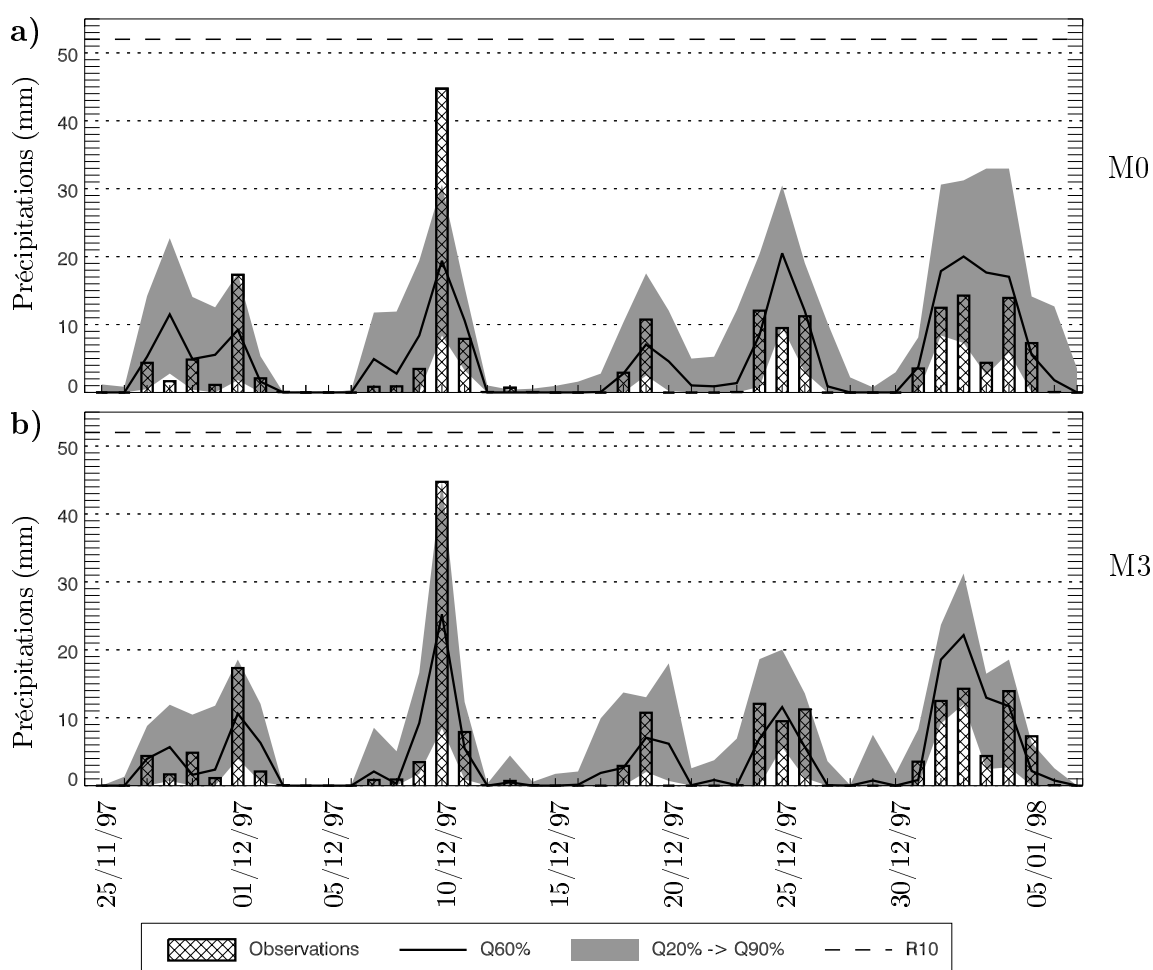


FIG. 7.8: Comparaison sur le groupement Arly, du 25/11/97 au 7/01/98, entre les précipitations observées et prévues par les algorithmes M0 (a)) et M3 (b)), en prévision parfaite.

Le deuxième exemple (figure 7.8), présenté sous une forme identique au précédent, concerne les prévisions émises du 25 novembre 1997 au 7 janvier 1998, sur le groupement **Arly**.

On constate sur cet exemple que les prévisions présentent les mêmes propriétés que précédemment, notamment en ce qui concerne leur bonne co-fluctuation avec les observations et le gain de finesse apporté par M3.

On note cependant que l'algorithme M0 montre une certaine faiblesse dans la prévision du fort cumul du 10 décembre 97, lequel dépasse les 44 mm alors que la pluie décennale sur ce groupement est de 52 mm. En effet, même si le quantile 90% atteint 30 mm, M0 émet par la suite des prévisions aussi fortes (25 décembre 97, 3 et 4 janvier 98), alors que les précipitations observées lors de ces journées-ci n'atteignent pas 15 mm. On est donc confronté à un défaut de discrimination de la part de M0.

De son côté, M3 parvient à faire un meilleur trié entre toutes ces journées. Les quantiles de prévisions sont ainsi fortement augmentés pour la journée du 10 décembre, alors qu'ils retrouvent des niveaux moyens là où M0 annonçait de fausses alertes.

7.3 Correction du biais ?

Une prévision probabiliste a vocation à être utilisée dans le cadre d'un problème de minimisation de l'espérance des coûts induits. Pour que cette utilisation soit rentable, il faut cependant que les prévisions ne soient pas biaisées, c'est à dire que globalement, les quantiles au non-dépassement $Q_{x\%}$ prévus ne doivent effectivement pas être dépassés dans exactement $x\%$ des cas.

Les *reliability charts*, décrits dans le chapitre 3 (cf. 3.1.3.3) permettent d'illustrer la présence éventuelle d'un biais à grande échelle dans les prévisions. La figure 7.9 présente ainsi celui associé à l'algorithme M1 sur l'ensemble de la période de calibration.

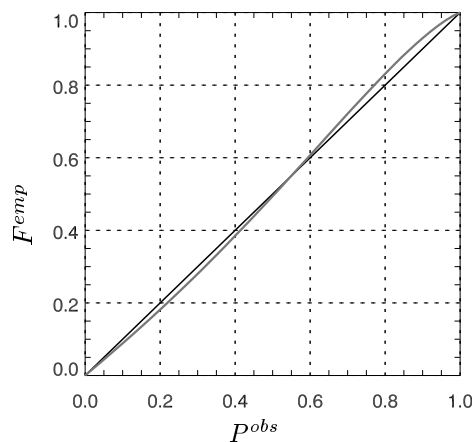


FIG. 7.9: Représentation du biais des prévisions (algorithme M1 sur période de calibration) : *reliability chart*.

On constate que les prévisions sont très peu biaisées, la fonction de répartition des quantiles observés s'écartant peu de la première bissectrice. Les faibles quantiles ont tendance à être globalement sous-estimés alors que les forts quantiles sont plutôt sur-estimés. Autrement dit, la dispersion des prévisions est un peu exagérée : le quantile $Q_{20\%}$ n'est pas atteint dans seulement 18% des cas, et le quantile $Q_{90\%}$ n'est dépassé que dans à peu près 7% des cas (au lieu de 10% théoriquement). Le biais est donc globalement faible et ne mérite pas d'être corrigé.

Evidemment, le biais n'est pas le même dans toutes les situations, et l'on sait que certaines sont fortement biaisées, comme celles présentant de très forts cumuls de précipitations. Il pourrait donc être intéressant de raffiner la description du biais en fonction de seuils de précipitations, voire en fonction de types de temps (le biais de prévision des situations anticycloniques estivales ne devant pas être le même que celui des entrées maritimes hivernales). Néanmoins, la pertinence de telles corrections dépend fortement de la fonction de coût employée par l'utilisateur final. Celui-ci donne-t-il en effet la même importance aux prévisions des forts cumuls qu'à celle de la simple occurrence des précipitations ?

Il appartient donc à chaque utilisateur d'aller éventuellement plus loin dans la description du biais, en fonction de ses besoins et de sa manière d'utiliser les prévisions.

7.4 Conclusions

A l'issue de la phase d'optimisation présentée dans la partie II, nous avons retenu 4 algorithmes afin de les valider sur un échantillon de situations cibles indépendant de celui utilisé pour la calibration. Ces algorithmes présentent des performances croissantes et représentent chacun un optimum en fonction de l'acceptation d'un certain degré de complexité :

- **M0** : algorithme de référence employé de façon opérationnelle (Guilbaud 1997),
- **M1** : idem M0, mais en utilisant les champs de géopotentiels $z_{1000+12}$ & z_{500+24} et une fenêtre d'analogie adaptée,
- **M2** : idem M1, mais en effectuant des regroupements régionaux des groupements pluviométrique et en utilisant des fenêtres d'analogies propres à chaque regroupement,
- **M3** : idem M2, mais en ajoutant une seconde sélection sur $PW\%$.

La validation de ces algorithmes a été effectuée avec succès, leur hiérarchie étant restée préservée. Les performances obtenues sur l'échantillon de validation ont même été supérieures à celles de la phase de calibration, prouvant ainsi qu'il n'y a pas eu de sur-optimisation.

Par ailleurs, l'utilisation de scores basés sur l'établissement de tables de contingence a permis d'avoir une vision plus détaillée des performances des algorithmes d'adaptation par analogie. Nous avons ainsi montré que :

- l'emploi de champs de géopotentiels à +12h permet d'améliorer la prévision des précipitations importantes,
- l'introduction de variables de second ordre telles que $PW\%$ autorise pour sa part des gains de performance pour la prévision des précipitations faibles,

- et ces améliorations se traduisent, un peu, par une augmentation de la capacité de détection des événements, mais surtout par une diminution du taux de fausses alertes.

Il est aussi apparu que nos algorithmes présentent un maximum d'efficacité pour la prévision des pluies faibles, plutôt que pour celle de l'occurrence ou non des précipitations. On assiste même à une chute des performances si l'on s'intéresse aux seuils de précipitations les plus élevés : le manque d'événements extrêmes au sein de l'archive pluviométrique handicape alors une utilisation quantitative. Néanmoins, d'un point de vue qualitatif, les prévisions demeurent tout à fait informatives.

Si l'utilisateur préfère manipuler une version déterministe des prévisions, il ne doit pas s'intéresser au même quantile en fonction du seuil de précipitations qui le préoccupe. Ainsi :

- $Q_{60\%}$ présente les meilleures performances en prévision d'occurrence de précipitations,
- $Q_{90\%}$ est le plus adapté à la prévision des précipitations moyennes à fortes,
- $Q_{20\%}$ permet de donner une limite inférieure aux prévisions, du fait d'un très faible taux de fausse alerte.

Sensibilité à l'archive météorologique



L'archive météorologique joue un rôle crucial dans une adaptation par analogie. De sa richesse et de son exactitude dépend effectivement la qualité des analogies que l'on sera en mesure de reconnaître.

Le choix de l'utilisation d'une archive plutôt qu'une autre peut donc se révéler décisif quant à l'utilité de l'adaptation. C'est pourquoi, afin de guider éventuellement un tel choix, nous allons nous attarder un peu dans ce chapitre sur la sensibilité des performances de l'adaptation aux divers paramètres de cette archive.

Nous allons commencer par nous interroger sur l'homogénéité réelle de nos archives : les réanalyses de l'archive NCEP/NCAR sont elles aussi homogènes que nous l'espérons et l'archive EDF aussi hétérogène que nous le redoutions ? Ce travail formera l'occasion de comparer les performances obtenues en utilisant l'une ou l'autre de ces archives et de mettre ainsi en évidence l'importance de l'homogénéité des données utilisées.

Dans un second temps, en restreignant la période sur laquelle on s'autorise à rechercher des situations analogues, nous essayerons d'identifier la longueur d'archivage nécessaire à l'obtention des meilleurs résultats. Enfin, nous aborderons la question de la résolution spatiale de l'archive : serait-il réellement utile d'utiliser des archives à haute résolution, ou pouvons-nous nous contenter de résolutions assez grossières, comme celle de l'archive LTHE (sous-ensemble des réanalyses NCEP/NCAR) ?

8.1 Homogénéité des archives météorologiques

Les deux archives météorologiques dont nous disposons, l'archive EDF et l'archive LTHE, ne présentent a priori pas la même homogénéité dans leurs données. En comparant les performances obtenues, selon que l'on emploie l'une ou l'autre, il devrait donc être possible de vérifier si cette homogénéité exerce une quelconque influence sur les résultats d'une adaptation par analogie.

8.1.1 Critique de l'homogénéité

L'une des principales motivations dans l'emploi des réanalyses NCEP/NCAR en tant qu'archive météorologique, outre l'importante quantité de variables atmosphériques disponibles, est l'espoir que ces données soient notablement plus homogènes que les archives constituées en compilant les analyses de modèles qui évoluent au cours du temps. Cependant, même si le modèle NWP utilisé pour produire des réanalyses est toujours identique, il n'en est pas de même pour les mesures disponibles pour initialiser ces réanalyses (cf. 2.1.3). Il est donc légitime de se demander si l'archive LTHE est aussi homogène que nous l'imaginions...

D'éventuelles ruptures de stationnarité peuvent par ailleurs être d'origines très variées. Citons par exemple les cas de figure suivants :

- modification du réseau de mesure,
- modification du système d'assimilation des données,
- ajout d'un nouveau type de données,
- anomalie temporaire de la circulation générale (phénomène de type El Niño),
- présence d'une dérive climatique...

8.1.1.1 Méthode de critique adoptée

Si nos archives étaient parfaitement homogènes, alors la probabilité de rencontrer une analogue à n'importe quelle situation donnée serait la même quelle que soit l'année à l'intérieur de laquelle on effectue la recherche¹. Ainsi, les meilleures analogues à un ensemble de situations cibles données devraient être équiréparties à l'intérieur de l'archive météorologique. Si ce n'est pas le cas, si elles se regroupent par exemple à un bout ou l'autre de l'archive, alors celle-ci n'est pas homogène.

C'est cette propriété que nous allons utiliser pour caractériser l'homogénéité de nos archives. Nous allons ainsi vérifier si l'on trouve autant de situations, analogues aux situations cibles de la période de calibration, tout au long de l'archive météorologique ou si celles-ci se rassemblent autour de certaines dates, selon la méthode que nous avons proposée (Bontron *et al.* 2002).

En pratique, nous procédons en découpant la période de calibration année par année. Pour chacune des 365 (ou 366) situations d'une année Y_j donnée, nous recherchons où se situent à l'intérieur de l'archive météorologique, année Y_j exclue, les 50 situations présentant le plus de similitude. Ces 18250 (ou 18300...) analogues se répartissent donc au sein de 43 années et nous pouvons faire le compte de leur nombre $N_{i,j}$ au sein d'une année Y_i donnée. Idéalement, si l'archive est parfaitement homogène, nous devons compter entre 424 et 426 analogues par année.

En recommençant l'opération pour toutes les années Y_j de la période de calibration, nous représentons le tableau complet des $N_{i,j}$ en fonction de l'année des situations cibles Y_j et de l'année des situations candidates Y_i . Un tel tableau nous indique à quel point une année est

¹Si nous voulions être parfaitement rigoureux, nous devrions plutôt parler de périodes de 2 ou 3 ans. En effet, une année précise peut être plus ou moins marquée par des phénomènes climatiques particuliers et ne pas présenter tous les types de situations possibles.

globalement analogue à une autre. Pour plus de commodité, nous représenterons ce tableau sous la forme d'une cartographie des $N_{i,j}$.

Cette méthode ne fait bien sûr ressortir qu'un seul aspect de l'homogénéité de nos archives : celle liée à la variable d'analogie choisie. Mais justement, c'est celle qui nous concerne dans l'utilisation des archives météorologiques.

8.1.1.2 Homogénéité de l'archive LTHE

Pour étudier l'homogénéité de l'archive LTHE (figure 8.1), nous avons considéré l'algorithme M0 (algorithme utilisé de façon opérationnelle), en retenant 50 situations analogues, afin d'être en mesure de comparer ultérieurement ces résultats avec ceux obtenus sur l'archive EDF par la DTG.

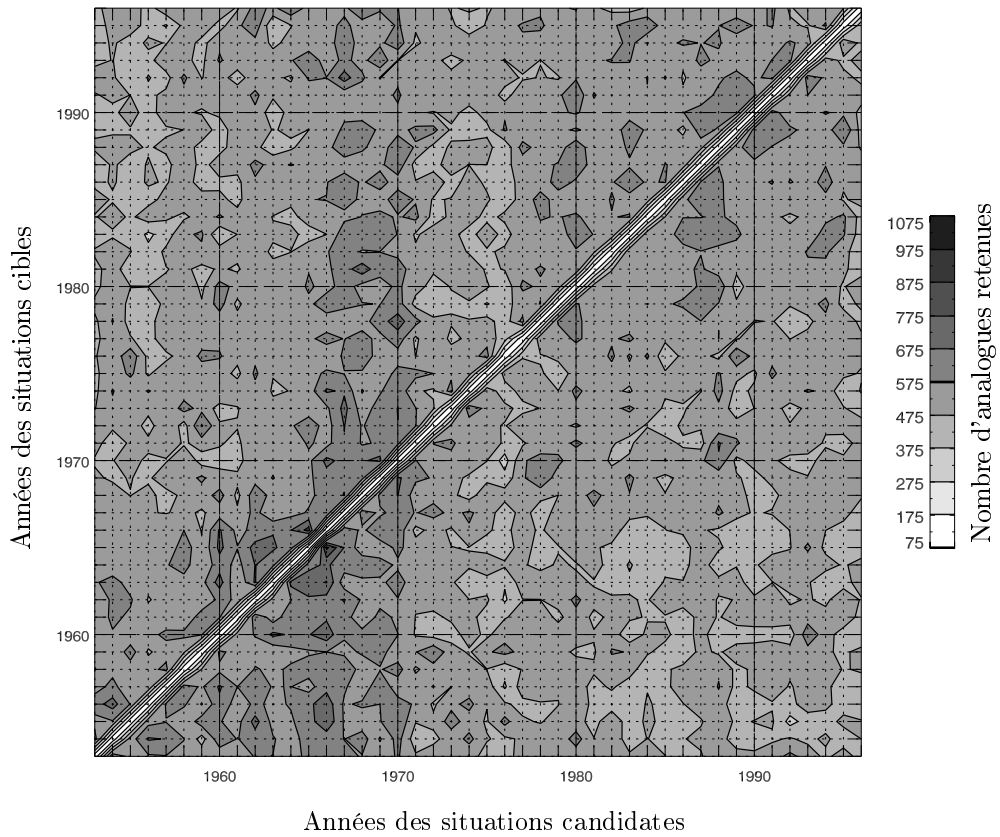


FIG. 8.1: Nombre d'analogues sélectionnées au sein de chaque année de l'archive LTHE en fonction de l'année de la situation cible.

A première vue, notre archive semble globalement homogène, les situations analogues étant sensiblement équiréparties au sein de l'archive. Néanmoins, si on regarde plus attentivement la figure 8.1, on constate que l'on peut différencier le comportement des situations cibles selon 3 périodes distinctes :

- ① Les années cibles allant de 1953 à 1967 privilégient fortement les situations candidates de la même période. Cet effet semble même plus marqué encore entre 1962 et 1967.

- ② Celles allant de 1967 à 1977 connaissent une répartition assez homogène de leurs analogues, avec néanmoins un pic de présence quasi systématique entre 1966 et 1971.
- ③ Les années cibles allant de 1977 à 1996 choisissent préférentiellement leurs analogues soit dans cette même période, soit à nouveau dans la période 1966-71.

On note donc deux périodes qui ne sont analogues qu'à elles-mêmes : 1962-68 et 1977-96. A ceci s'ajoute de façon systématique, quelle que soit l'année cible, la sous-représentation des années candidates 1971-77 et la sur-représentation des années 1967-71, ce qui représente en tout quatre ruptures de stationnarités : 1962, 1967, 1971 et 1977.

L'archive LTHER, basée sur les réanalyses NCEP/NCAR n'est donc pas complètement homogène. Isabelle Pocard (2000) a étudié en détail ces réanalyses et elle y a décelé elle aussi quatre ruptures de stationnarité, qui ne correspondent pas toujours à celles que nous mettons en évidence ici :

- Une rupture majeure en 1967-68, que nous retrouvons nous aussi. D'après I. Pocard, elle s'explique par la nette augmentation, autour de cette époque, de la densité des observations (données sols et radiosondages) assimilées par les réanalyses. Nous sommes donc ici dans une période transitoire.
- Et trois ruptures mineures en 1976-77, 1983-84 et 1987-88 (nous ne retrouvons ici que la trace de la première). Cet auteur indique qu'elles semblent liées pour leur part à des problèmes d'assimilation des données satellitales ainsi qu'aux nombreux changements de satellites intervenus au cours des années 80.

L'hétérogénéité de l'archive LTHER est donc a priori d'origine artificielle : elle ne correspond pas à des ruptures climatiques réelles. Il n'y a ainsi aucun changement climatique sensible dans nos 44 années de données de géopotentiels.

8.1.1.3 Homogénéité de l'archive EDF

A la suite de la critique de l'archive LTHER, une étude similaire a été menée sur l'archive EDF (figure 8.2, Laplanche 2002, communication personnelle), à partir de l'algorithme d'analogie implanté opérationnellement à la DTG (M0).

L'archive EDF présente des hétérogénéités bien plus prononcées que celles de l'archive LTHER. On note ainsi que la période qui va de 1975 à 1985 se démarque notablement du reste de l'archive. Les situations cibles qui appartiennent à cette période reconnaissent en effet principalement leurs analogues dans ces mêmes années. Les périodes 1953-55, 1956-59 et 1985-97 présentent par ailleurs des comportements similaires, mais de façon un peu moins marquée.

On constate donc la présence de trois ruptures dans cette archive : 1955, 1975 et 1985. En se tournant vers l'histoire de la formation de cette archive, Maud Laplanche a pu expliquer deux de ces ruptures :

- La rupture de 1975 correspond au passage de données directement mesurées par des stations de radiosondage, à des données analysées par un modèle NWP, puis interpolées aux points

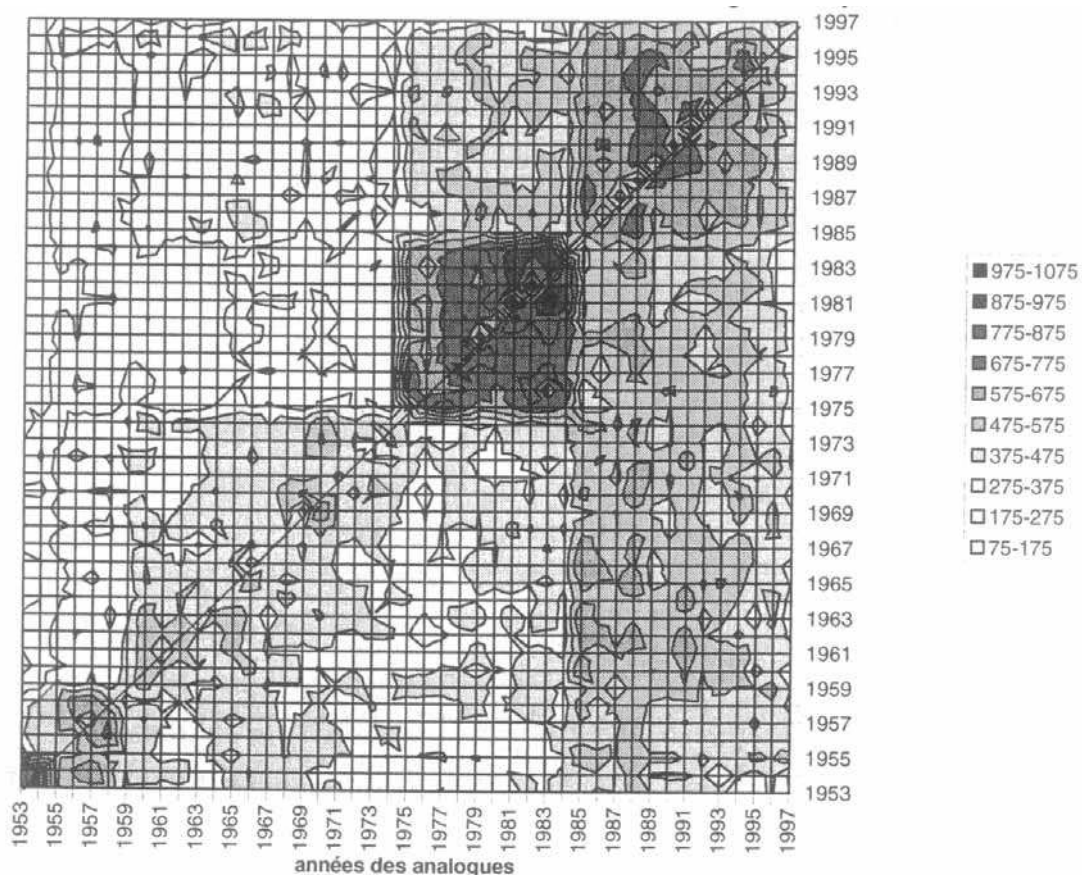


FIG. 8.2: Nombre d'analogues sélectionnés au sein de chaque année de l'archive EDF en fonction de l'année de la situation cible (Document EDF).

de radiosondage. La texture des données n'est alors plus la même : une rapide étude sur les gradients de pression de l'archive a montré que les données analysées étaient plus « lisses » que les données mesurées.

- La rupture de 1985 s'explique par le changement du modèle NWP utilisé pour l'analyse : le modèle EMERAUDE a été abandonné au profit du modèle ARPEGE. Outre des modifications dans la physique du modèle utilisé, qui interviennent toutefois peu sur l'analyse des champs de pression, ce changement a apporté une importante augmentation de la résolution de l'analyse. Nous sommes ainsi passé d'une résolution horizontale de 200 km pour EMERAUDE, à une résolution de l'ordre de 20 km pour ARPEGE. Ce changement se traduit encore une fois par une modification de la texture des champs archivés, mais aussi de leur qualité.

Les hétérogénéités de l'archive EDF sont donc elles aussi d'origines artificielles. L'absence de points communs manifestes dans les répartitions obtenues pour les deux archives tend par ailleurs à confirmer leur caractère artificiel.

Finalement, l'archive EDF apparaît comme nettement plus hétérogène que l'archive LTHE. Les réanalyses NCEP/NCAR présentent donc bien un gain d'homogénéité, même si des ruptures de stationnarité demeurent. On peut donc espérer que les nouvelles générations de réanalyses présenteront une homogénéité encore supérieure.

8.1.2 Relation entre homogénéité et performance

Les deux archives dont nous disposons ne présentent pas le même degré d'homogénéité. Nous allons donc comparer les performances que l'on peut obtenir en utilisant soit l'une, soit l'autre. Néanmoins, pour que cette comparaison soit significative, il faut bien sûr que les algorithmes d'adaptation soient exactement les mêmes. Or, ces deux archives ne contiennent pas les mêmes données et n'ont pas non plus la même résolution (cf. 2.1). Il nous faut donc au préalable définir un protocole commun.

8.1.2.1 Algorithme de comparaison

Les résolutions spatiales des deux archives sont incompatibles entre elles (figure 8.3), c'est pourquoi il est nécessaire de les ré-exprimer selon une grille commune. Pour ce faire, nous avons décidé d'utiliser la grille déjà associée à l'archive LTHE : grille régulière en longitude et latitude et possédant une maille de $2.5 \times 2.5^\circ$. En effet, l'interpolation des données de l'archive LTHE sur une grille de plus grande résolution, sans ajout d'information, ne leur conférerait pas pour autant une structure comparable à celle des données de l'archive EDF : à l'effet de l'hétérogénéité s'ajouterait alors celui de la résolution. A l'inverse, l'expression des champs de l'archive EDF sur une grille de $2.5 \times 2.5^\circ$, de résolution plus grossière, leur donne effectivement une structure comparable à celle des données contenues dans l'archive LTHE.

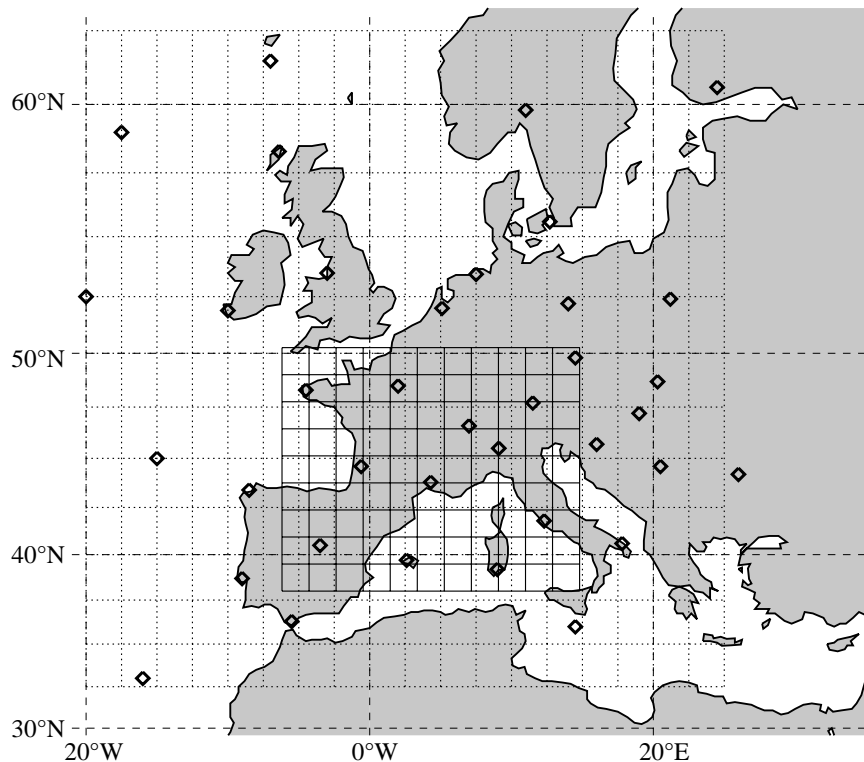


FIG. 8.3: Emplacement des stations de radiosondage utilisées pour constituer l'archive EDF (losange), et définition des grilles de l'archive EDF (traits pleins) ainsi que de l'archive LTTH (traits pointillés).

Une nouvelle archive, nommée EDF_{2.5}, a donc été créée en interpolant les points de radiosondage constitutifs de l'archive EDF sur la grille 2.5x2.5° de l'archive LTHE. Cette interpolation est effectuée par une méthode spline identique à celle utilisée pour la création de l'archive EDF originelle à partir des points de radiosondage.

Les archives EDF_{2.5} et EDF ne contenant pas de variables autres que les géopotentiels, la comparaison ne peut s'effectuer que sur une sélection de premier ordre. Nous avons donc choisi d'utiliser la méthode de référence M0, employée de façon opérationnelle par EDF. Notons toutefois que selon l'archive considérée, la fenêtre d'analogie diffère légèrement :

- fenêtre opérationnelle pour l'archive EDF :
lon0 = -6.2°E, lat0 = 38°N, dlon = 19.1, dlat = 12.3,
- fenêtre la plus proche possible de la fenêtre opérationnelle pour les autres archives :
lon0 = -7.5°E, lat0 = 37.5°N, dlon = 20°, dlat = 12.5°.

8.1.2.2 Performances globales

Le tableau 8.1 présente les performances moyennes sur l'ensemble des groupements pluviométriques (en terme de **Comp. CRPS**) ainsi obtenues sur la période de calibration, en fonction de l'archive décrivant soit les situations cibles, soit les situations candidates.

		SITUATIONS CIBLES		
		EDF	EDF _{2.5}	LTHE
SITUATIONS CANDIDATES	EDF	36.82%	-	-
	EDF _{2.5}	-	36,24%	36,31%
	LTHE	-	35,79%	37,27%

TAB. 8.1: Performances (**Comp. CRPS**) obtenues en fonction de l'archive dont sont issues les situations candidates et les situations cibles.

Comme nous l'espérons du fait de sa plus grande homogénéité, l'utilisation de l'archive LTHE pour caractériser à la fois les situations cibles et les situations candidates, apporte un gain de performance par rapport à l'utilisation de l'archive EDF_{2.5}. Toutefois, ce gain n'est que de 2.84%, ce qui reste finalement assez modeste.

En outre, lorsque l'on considère deux archives différentes pour décrire les situations cibles et les situations candidates, il s'avère que les performances ont tendance à diminuer par rapport à l'utilisation d'une archive unique. Donc, en plus d'une homogénéité des archives dans le temps, il est préférable de respecter une certaine homogénéité entre les données de l'archive historique et celles qui décrivent les situations cibles.

Dans le cadre d'une utilisation opérationnelle, il est donc intéressant de disposer d'une archive météorologique analysée avec un modèle NWP comparable à celui utilisé pour prévoir les situations cibles.

Le gain de performance apporté par la meilleure homogénéité de l'archive LTHE n'est donc pas suffisant pour justifier à lui seul le remplacement de l'archive opérationnelle EDF, le modèle

NWP utilisé n'étant pas forcément homogène avec les réanalyses NCEP/NCAR. Le principal intérêt de cette archive réside plutôt dans la richesse de l'information fournie et la possibilité de choisir parmi plus de variables et d'échéances. Ce type d'archive autorise ainsi la mise en œuvre d'analogies reposant sur deux sélections successives, ce qui permet une augmentation significative des performances, comme nous l'avons vu dans les chapitres précédents.

Le tableau 8.1 montre par ailleurs que la perte de résolution due au passage de l'archive EDF à l'archive EDF_{2.5} engendre une légère perte de performance. Ce résultat souligne un autre inconvénient des réanalyses : leur résolution assez lâche. Nous reviendrons plus loin dans ce chapitre (cf. 8.3) sur l'impact de la résolution de l'archive.

8.1.2.3 Performances annuelles

Les écarts de performances entre les deux archives EDF_{2.5} et LTHE ne sont pas constants tout au long de la période de calibration, comme le montre la figure 8.4. Selon l'année des situations cibles, l'archive EDF_{2.5} peut même se montrer plus utile que l'archive LTHE.

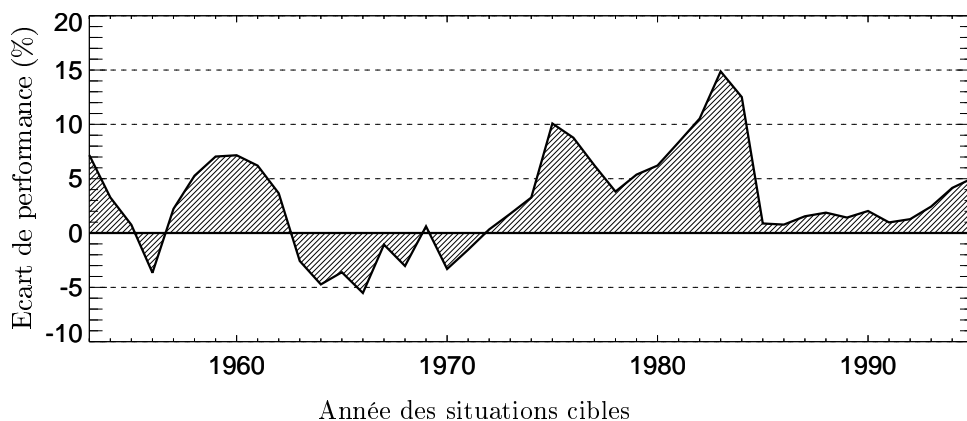


FIG. 8.4: Ecarts relatifs de performances (**Comp. CRPS**) entre les archives EDF_{2.5} et LTHE, en fonction de l'année des situations cibles.

La principale période pour laquelle les écarts de performance sont en faveur de l'archive EDF_{2.5}, prend place de 1963 à 1971. Or, ceci coïncide avec la période fortement hétérogène de l'archive LTHE, pour laquelle les situations cibles recherchent de façon effective leurs analogues dans un panel limité de situations candidates. Cette impossibilité de considérer des situations candidates dans l'intégralité de l'archive, du fait de l'hétérogénéité de cette période, nuit aux performances de l'archive LTHE. Dans le même temps, la période 1963-71 est relativement homogène en ce qui concerne l'archive EDF : le panel des situations candidates possibles étant plus riche, l'adaptation est de meilleure qualité.

A l'inverse, de 1975 à 1984, l'archive LTHE est nettement plus utile que l'archive EDF_{2.5}. Or, maintenant, cette période est associée à une forte hétérogénéité de l'archive EDF, alors que l'archive LTHE s'y présente comme relativement homogène. Dès 1985, alors que l'archive EDF voit une amélioration de son homogénéité, l'écart de performance entre les deux archives diminue.

Il apparaît donc que l'homogénéité des archives joue un rôle important dans la qualité d'une adaptation par analogie. On trouve ainsi des écarts de performances qui peuvent atteindre 15 % sur l'année. En pratique, les hétérogénéités limitent le choix des situations cibles : tout se passe comme si l'archive était plus courte qu'elle ne l'est réellement. Il est donc possible d'améliorer les performances de l'adaptation par analogie, uniquement en disposant d'une archive plus homogène, à condition toutefois que la description des situations cibles soit elle aussi homogène avec l'archive.

8.2 Longueur d'archive nécessaire

Comme nous venons de le voir, l'influence de l'homogénéité des archives météorologiques se ramène à celle de la longueur de la période couverte. Finalement, si on a le choix entre deux archives météorologiques, vaut-il mieux privilégier une archive longue, mais potentiellement hétérogène, ou une archive plus courte, mais garantissant une meilleure homogénéité ?

Nous allons nous intéresser ici à l'effet de la longueur de l'archive météorologique LTHE sur les performances obtenues. Pour ce faire, nous limitons volontairement les situations candidates (issues de la période de calibration, 1953-96) par tranches de 5 ans à partir de 1953. Ainsi, nous considérerons successivement les situations candidates de 1953 à 1996, puis celles allant de 1958 à 1996, et ainsi de suite jusqu'à ne plus utiliser que celles comprises entre 1993 et 1996. En ce qui concerne les situations cibles, celles-ci parcourront toujours l'intégralité de la période de validation (1997-2001).

En premier lieu, étant donné que notre archive n'est pas parfaitement homogène, il convient de s'intéresser à la représentativité de chacune des tranches de 5 années des situations candidates, vis à vis des situations cibles. La figure 8.5 présente la répartition des situations reconnues comme analogues au sein de ces différentes sous-périodes, en utilisant l'algorithme M1 (cf. 7.1.2).

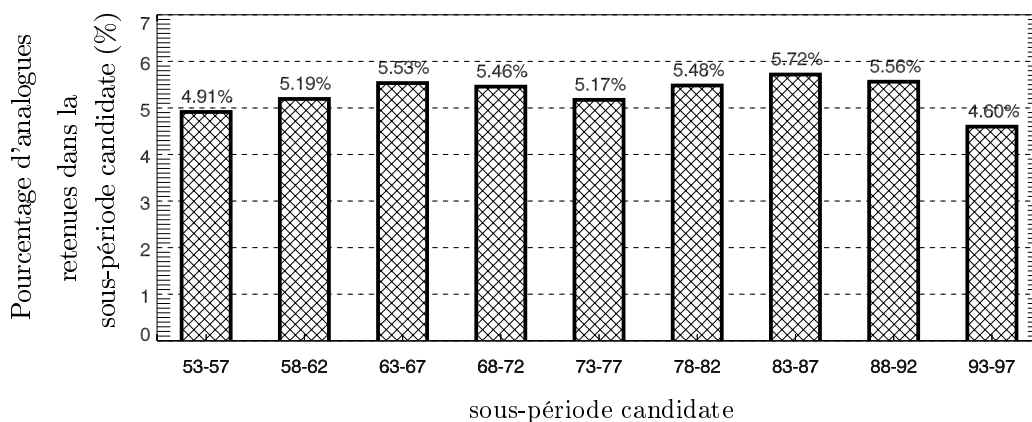


FIG. 8.5: Répartition des analogues au sein de sous-périodes candidates de l'archive LTHER pour la période cible 1997-2001.

Cette répartition apparaît comme relativement homogène, la part de chacune des sous-périodes variant de 4.91 % à 5.72 %. On note que la sous-période 1973-77 représente un minimum relatif : nous retrouvons ici, mais de façon atténuée, la sous-représentation systématique des situations candidates de ces années que mettait en évidence la figure 8.1. La sous-période 1993-97 ne représente par ailleurs que 4.60 % des analogues sélectionnées, mais ceci est dû au fait qu'elle ne fait que 4 ans : après correction, la part de cette sous-période est de 5.70 %. Ces résultats nous permettent de considérer que chacune des sous-périodes est équivalente au niveau de la quantité d'information apportée aux algorithmes d'adaptation. Nous n'avons de ce fait pas de précautions particulières à prendre pour tester la sensibilité de nos algorithmes à la longueur de l'archive.

Comme l'illustre la figure 8.6, si l'on considère l'ensemble des situations cibles de la période de validation, l'augmentation des performances globales ralentit fortement au delà de 30 ans d'archives, que ce soit pour une analogie basée sur une seule sélection (M1) ou sur deux sélections (M3). L'utilisation d'une archive plus longue apporte peu d'améliorations dans la prévision des situations courantes.

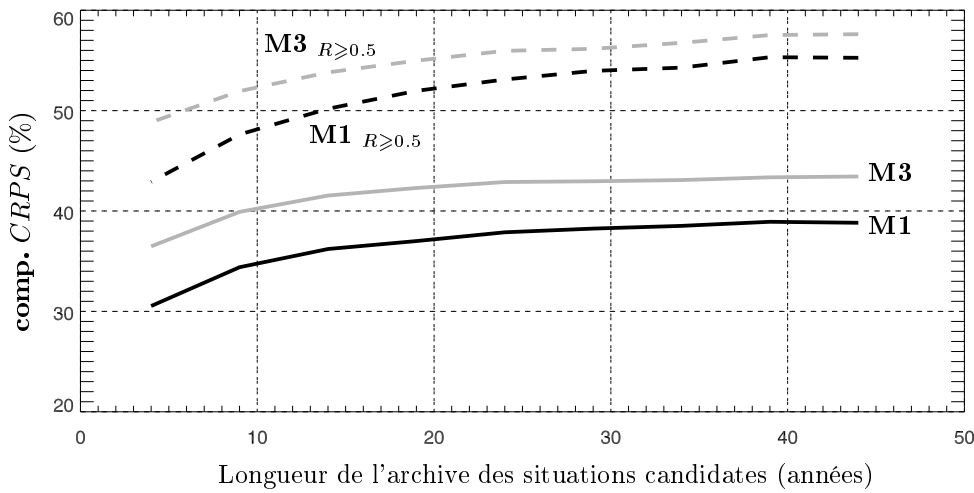


FIG. 8.6: Performance globale sur l'ensemble des groupements des algorithmes M1 et M3 sur la période de validation, exprimée en terme de **Comp. CRPS**, en fonction de la longueur de l'archive des situations candidates. **Courbes pleines** : ensemble des situations cibles de la période de validation ; **courbes pointillées** : situations cibles telles que $R \geq 0.5$. ($R = \sqrt{\frac{R_{\text{cible}}}{R_{10}}}$)

Toutefois, si l'on s'intéresse aux seules situations cibles présentant de fortes précipitations (courbes en pointillés sur la figure 8.6), on constate que même avec 40 ans d'archive nous n'observons pas de plafonnement des performances. Les situations exceptionnelles bénéficient donc de l'utilisation d'un panel plus large de situations candidates. Elles sont toutefois peu nombreuses, par définition, ce qui explique le peu d'impact de ce gain sur les performances globales passé une trentaine d'années d'archive.

Les résultats quantitatifs apportés ici permettent de conclure quant aux raisons de l'amélioration des performances observées en validation par rapport à la phase de calibration. La sensibilité de l'adaptation à la longueur de l'archive météorologique montre en effet que l'utilisation d'une année d'archive supplémentaire (passage de 43 à 44 ans) est insuffisante pour expliquer cette

amélioration. Par contre, la relative homogénéité de l'archive vis à vis de la période de validation représente, par rapport à des situations cibles allant de 1953 à 1968, un gain de l'ordre de 15 années d'archive effective. Ce sont donc les problèmes d'homogénéité de l'archive qui expliquent cette amélioration de performance.

Il est en outre informatif de s'intéresser au nombre optimal de situations analogues à retenir, en fonction de la longueur de l'archive. Ainsi, la figure 8.7 montre que l'emploi d'archives courtes impose une nette diminution de cet optimum. En effet, l'éventail de situations candidates étant restreint, il est plus difficile d'en trouver un grand nombre présentant de fortes similitudes avec la situation cible.

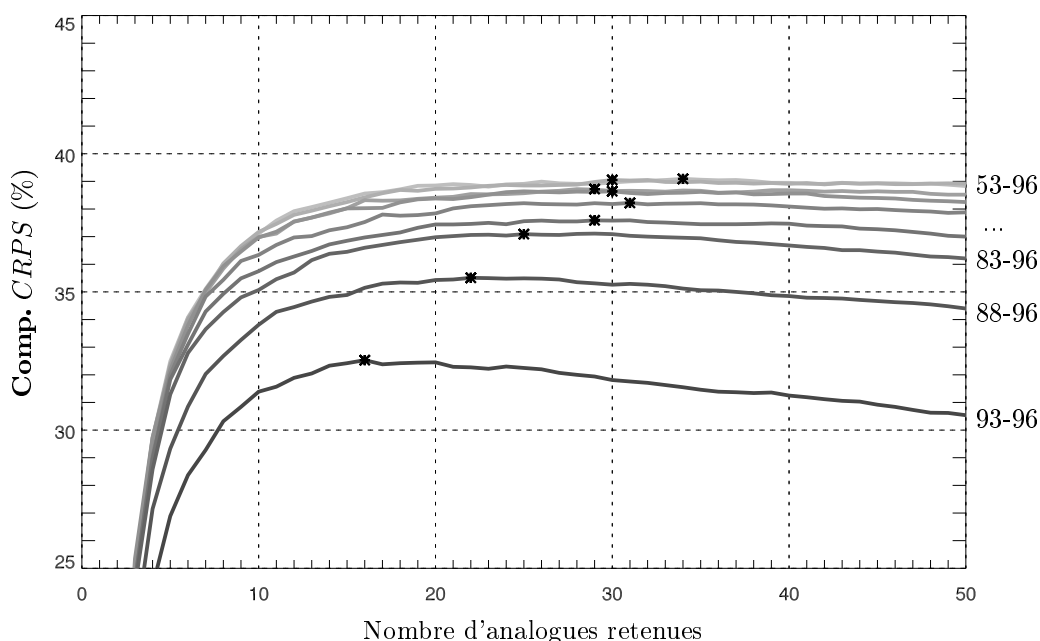


FIG. 8.7: Evolution des performances globales sur l'ensemble des groupements de l'algorithme M1 sur la période de validation, en fonction du nombre d'analogues retenues, pour différentes longueurs d'archives de situations candidates. Les étoiles représentent les maxima de performances de chaque courbe.

Néanmoins, au delà d'une vingtaine d'année de données, parallèlement au plafonnement des performances, le nombre optimal de situations analogues semble se stabiliser autour de 30. L'augmentation du nombre d'analogues sélectionnées, même si l'on dispose d'une grande quantité d'analogues de bonne qualité, n'entraîne qu'un sur-échantillonnage des fonctions de répartition conditionnelles des précipitations, sans augmentation des performances.

8.3 Effet de la résolution spatiale

Pour quantifier la sensibilité de nos algorithmes à la résolution de l'archive, il suffit de faire tourner le même algorithme avec différentes résolutions et de comparer les performances entre elles. Malheureusement, nous ne disposons pas d'archives météorologiques de plus fine résolution,

telles que les réanalyses européennes ERA-40. Notre seule possibilité est donc de dégrader de plus en plus fortement l'archive dont nous disposons. Nous allons ainsi tester l'utilisation de l'archive LTHE en considérant successivement :

- l'ensemble des points de grille ($2.5 \times 2.5^\circ$),
- un point sur deux ($5 \times 5^\circ$),
- un point sur trois ($7.5 \times 7.5^\circ$)
- et enfin un point sur six ($15 \times 15^\circ$).

Bien sûr, pour que les résultats soient comparables, il faut que l'extension des fenêtres d'analogie utilisées soit la même pour chacune des résolutions testées. Nous allons donc considérer un algorithme M1bis similaire à M1 (une seule sélection, fenêtre commune à l'ensemble des groupements), mais dont la fenêtre d'analogie est compatible avec ces quatre résolutions (figure 8.8). La fenêtre ainsi choisie est légèrement plus grande que la fenêtre optimale : $\text{lon}0 = -10^\circ\text{E}$, $\text{lat}0 = 35^\circ\text{N}$, $\text{dlon} = 30^\circ$ et $\text{dlat} = 15^\circ$.

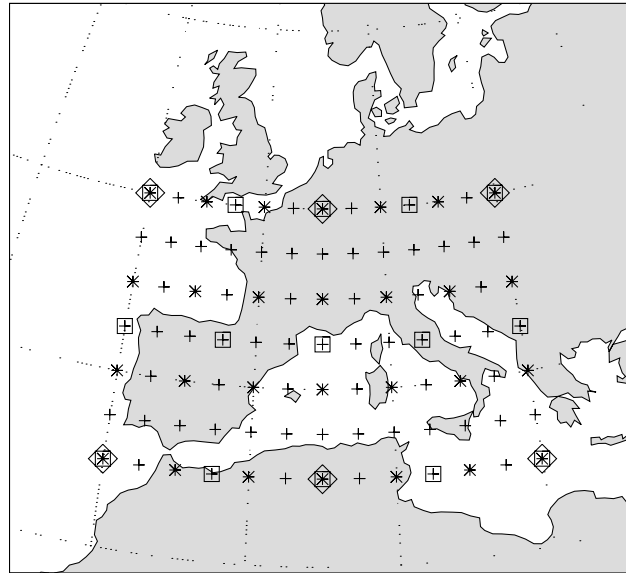


FIG. 8.8: Fenêtre d'analogie utilisée avec les différentes résolution : $2.5 \times 2.5^\circ$ (croix), $5 \times 5^\circ$ (étoiles), $7.5 \times 7.5^\circ$ (carrés) et $15 \times 15^\circ$ (losanges).

L'évolution des performances de cet algorithme (figure 8.9, courbe noire), montre que l'augmentation de la résolution entraîne peu de gain sur les performances en deçà d'une maille de 5° . La résolution des réanalyses NCEP/NCAR semble suffisante pour rechercher des analogies sur la circulation synoptique. La portée d'une variable telle que la hauteur du géopotentiel HGT est en effet plutôt de l'ordre de 20° . En outre, les géopotentiels sont des champs peu bruités. Un échantillonnage effectué à $2.5 \times 2.5^\circ$ est donc suffisant pour capter l'essentiel de l'information. Il est donc peu probable que l'emploi d'archives possédant une résolution de $1 \times 1^\circ$ génère une hausse vraiment sensible des performances.

Par ailleurs, nous avons introduit un algorithme M3bis comportant une analogie de second ordre similaire à celle de M3, basée sur la variable de second ordre $\text{PW}\%$. La première sélection de

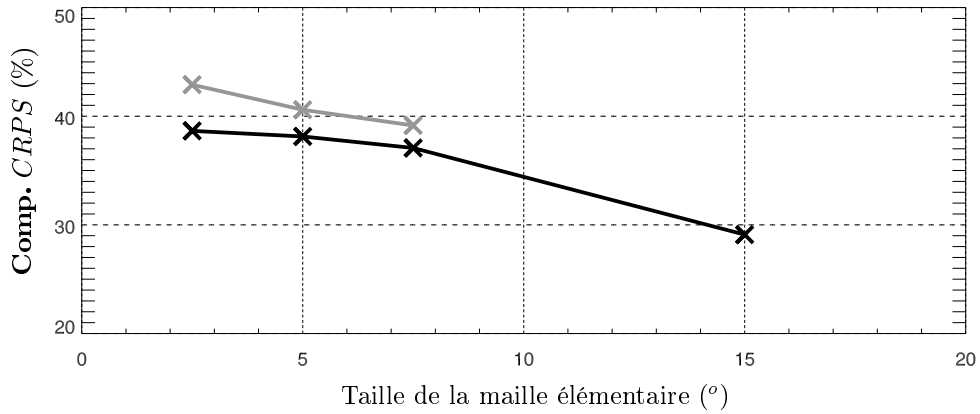


FIG. 8.9: Evolution des performances globales sur l'ensemble des groupements en fonction de la taille de la maille élémentaire de l'archive météorologique. **Courbe noire** : une seule sélection ; **courbe grisée** : deux sélections.

M3bis est identique à M1bis, à la différence que le premier panier d'analogie $N1$ est maintenant égal à 70 (au lieu de 50). Pour toutes les résolutions, les fenêtres d'analogie de second ordre sont constituées par les 4 points de grille les plus proches de chaque regroupement. Nous ne sommes donc plus exactement dans le même cadre que précédemment, l'extension de ces fenêtres étant ici dépendante de la résolution. Cette approche est toutefois celle qui nous semble la plus pertinente pour des variables locales telles que PW%, étant donné que l'on ne cherche pas ici à reconnaître une forme, mais plutôt à estimer un niveau moyen sur une zone restreinte autour des groupements cibles.

On constate (figure 8.9, courbe grisée) que les changements de résolution n'ont pas le même impact sur M3bis : le gain apporté par l'analogie de second ordre est d'autant plus sensible que la résolution est fine. De plus, l'évolution des performances ne semble pas exhiber de comportement asymptotique lorsque la résolution augmente, comme cela était le cas avec M1bis. Le caractère très local de l'analogie de second ordre ressort ici pleinement.

Les réanalyses à $2.5 \times 2.5^\circ$ sont donc fort certainement sous-optimales pour l'utilisation de M3 et M3bis. L'emploi d'archives plus fines présente ainsi un fort potentiel d'amélioration pour ces algorithmes d'analogie en deux temps.

8.4 Conclusions

Nous avons étudié au cours de ce chapitre la sensibilité de nos algorithmes d'analogie aux différents paramètres de l'archive météorologique.

Dans un premier temps, nous nous sommes attachés à vérifier l'homogénéité des archives dont nous disposons. Il est ainsi apparu que même si l'archive LTRE, issue des réanalyses NCEP/NCAR, est effectivement plus homogène que l'archive EDF, elle n'en présente pas moins certaines hétérogénéités. Or, ces hétérogénéités sont préjudiciables aux performances d'une adap-

tation par analogie, car elles limitent de façon artificielle la quantité de situations candidates exploitables.

Dans un second temps, nous avons testé l'impact de la profondeur des archives utilisées. Nous avons bien sûr montré que les performances croissent avec la longueur de l'archive. Cependant, ce sont surtout les situations engendrant de forts cumuls qui bénéficient de l'allongement de l'archive, les situations plus courantes se contentant d'une trentaine d'années de profondeur.

Enfin, nous avons voulu tester la sensibilité des algorithmes d'analogie à la résolution de l'archive. Or, il se trouve que cette sensibilité est différente selon que l'on considère un algorithme d'analogie comportant une ou deux étapes de sélection :

- la première sélection, basée sur la circulation synoptique via les géopotentiels HGT, montre un plafonnement de ses performances pour des mailles de taille inférieure à 5° ,
- alors que la seconde sélection voit ses performances nettement augmenter avec la résolution.

Le caractère local de la seconde analogie est ainsi pleinement démontré. Ces résultats montrent par ailleurs que les réanalyses NCEP/NCAR possèdent sûrement une résolution sous-optimale pour ces algorithmes à deux sélections. Dans ce contexte, l'utilisation des réanalyses du centre européen ERA-40, à $1 \times 1^\circ$ de résolution pourrait être un facteur de gain non négligeable.

9

Utilisation opérationnelle : étude de cas

« Ceux qui ont contemplé cette catastrophe des lieux élevés où elle les a surpris ont vu flotter pêle-mêle, ici de riches débris et de précieux meubles, là des ustensiles de campagne, ici des charrues, des bœufs avec leurs bouviers, là des troupeaux abandonnés à eux-mêmes et, au milieu de tout cela, des troncs d'arbres et des poutres arrachées des maisons. Mais les lieux mêmes où le fleuve n'est pas parvenu n'ont pas pour autant été épargnés par le désastre. Au lieu de l'inondation, ils connurent une pluie continuelle, des trombes d'eau s'échappant des nuages ; les murs qui entouraient ces riches domaines furent emportés, les monuments mêmes ébranlés et renversés. »

Pline le Jeune, Lettres (VIII, 17), traduction de Yves Hucher.

L'une des principales applications de l'adaptation par analogie est bien sûr la prévision opérationnelle des précipitations. Cependant, tout le travail effectué jusqu'ici a été mené dans un contexte de prévision parfaite. C'est à dire que les variables d'analogie qui servent à décrire les situations cibles étaient directement mesurées (ou plutôt *analysées*, pour être exact) : l'incertitude sur ces données était donc aussi minime que possible. Dans ce contexte, les performances atteintes en explication des pluies étaient uniquement le fruit des algorithmes d'adaptation.

Evidemment, dans un cadre de prévision opérationnelle, l'analyse de la situation cible n'est pas disponible au moment où l'on doit émettre une prévision : la description de cette situation se fait alors via un modèle NWP. Les données sur lesquelles se base la recherche d'analogie sont donc déjà entachées d'une certaine incertitude, laquelle vient s'ajouter à l'incertitude intrinsèque de l'algorithme d'adaptation. Les performances opérationnelles peuvent ainsi s'éloigner de celles présentées dans les chapitres précédents.

C'est pourquoi nous allons étudier dans ce dernier chapitre quelques exemples d'applications de nos algorithmes en prévision réelle. Pour ce faire, nous avons dans un premier temps dû mettre en œuvre au LTHE un système opérationnel qui permette de disposer tous les jours des prévisions d'un modèle NWP. Grâce à l'archivage de ces prévisions, il nous a ensuite été possible de tester différents algorithmes dans un contexte opérationnel. L'évènement pluviométrique intense survenu dans les Cévennes les 8 et 9 septembre 2002 fournira ainsi l'occasion d'une étude de cas.

9.1 Les prévisions opérationnelles au LTHE

9.1.1 Le dispositif opérationnel prototype

Les premiers ingrédients nécessaires à une application opérationnelle de nos algorithmes sont les sorties d'un modèle météorologique NWP, afin d'obtenir une prévision des variables d'analogie des situations cibles. Pour avoir accès tous les jours à ces données, nous avons mis en place au LTHE un dispositif automatisé. Il tourne sous Linux et se compose d'une chaîne opérationnelle, telle que schématisée par la figure 9.1. Le modèle retenu est le modèle GFS du NCEP (anciennement connu sous le nom AVN).

En premier lieu, à 6h heure locale, un premier **scripte de commande**, chargé de préparer l'initialisation de la chaîne, lance successivement les opérations suivantes :

- ① Lecture de la date de la journée en cours sur l'horloge interne du réseau informatique et préparation des fichiers d'initialisation (fichier `.ini`), qui seront lus par les différents programmes activés ultérieurement.
- ② Préparation de l'arborescence informatique (création de nouveaux répertoires) afin de faciliter le stockage des résultats de la prévision en cas de changement de mois.

Ensuite, à 10h heure locale, un second **scripte de commande**, qui rassemble les opérations de prévision proprement dite, entreprend dans l'ordre les actions suivantes :

- ① Connexion via **FTP** au serveur qui rassemble les sorties du modèle GFS et rapatriement des échéances nécessaires (de +0h à +168h par pas de 12h) sous forme de fichiers au format **Grib**.
- ② Extraction à partir des fichiers **Grib** des variables d'intérêt. Ces données, disponibles originellement sur l'ensemble du globe avec une résolution de $1\times 1^\circ$, sont interpolées sur un domaine restreint avec une résolution $2.5\times 2.5^\circ$ (grille de l'archive LTHE), avant d'être archivées au format **NetCDF**¹.
- ③ Mise en œuvre des algorithmes d'analogie à partir des données mises au format **NetCDF**. Les résultats de ces prévisions sont écrits dans des fichiers **ASCII** qui sont ensuite stockés dans une arborescence particulière. Pour chaque jour d'émission de la prévision, chacun des algorithmes d'analogie utilisés produit un fichier de résultat différent.

La visualisation de ces prévisions peut ensuite s'effectuer à n'importe quel moment via une interface graphique que nous avons développée sous **IDL**, et dont le mode d'emploi est reproduit en annexe A. Cette interface interprète uniquement les fichiers **ASCII** de résultats, tous les calculs de prévisions étant effectués une fois pour toute par la chaîne opérationnelle. Elle permet entre autres de visualiser les réanalyses NCEP/NCAR correspondantes aux situations analogues et de les comparer aux prévisions GFS de la situation cible. Etant donné que ces fichiers sont archivés, notre interface peut aussi servir à visualiser d'anciennes prévisions.

¹Les variables et le domaine considérés correspondent exactement au contenu de l'archive météorologique LTHE, décrite au chapitre 2

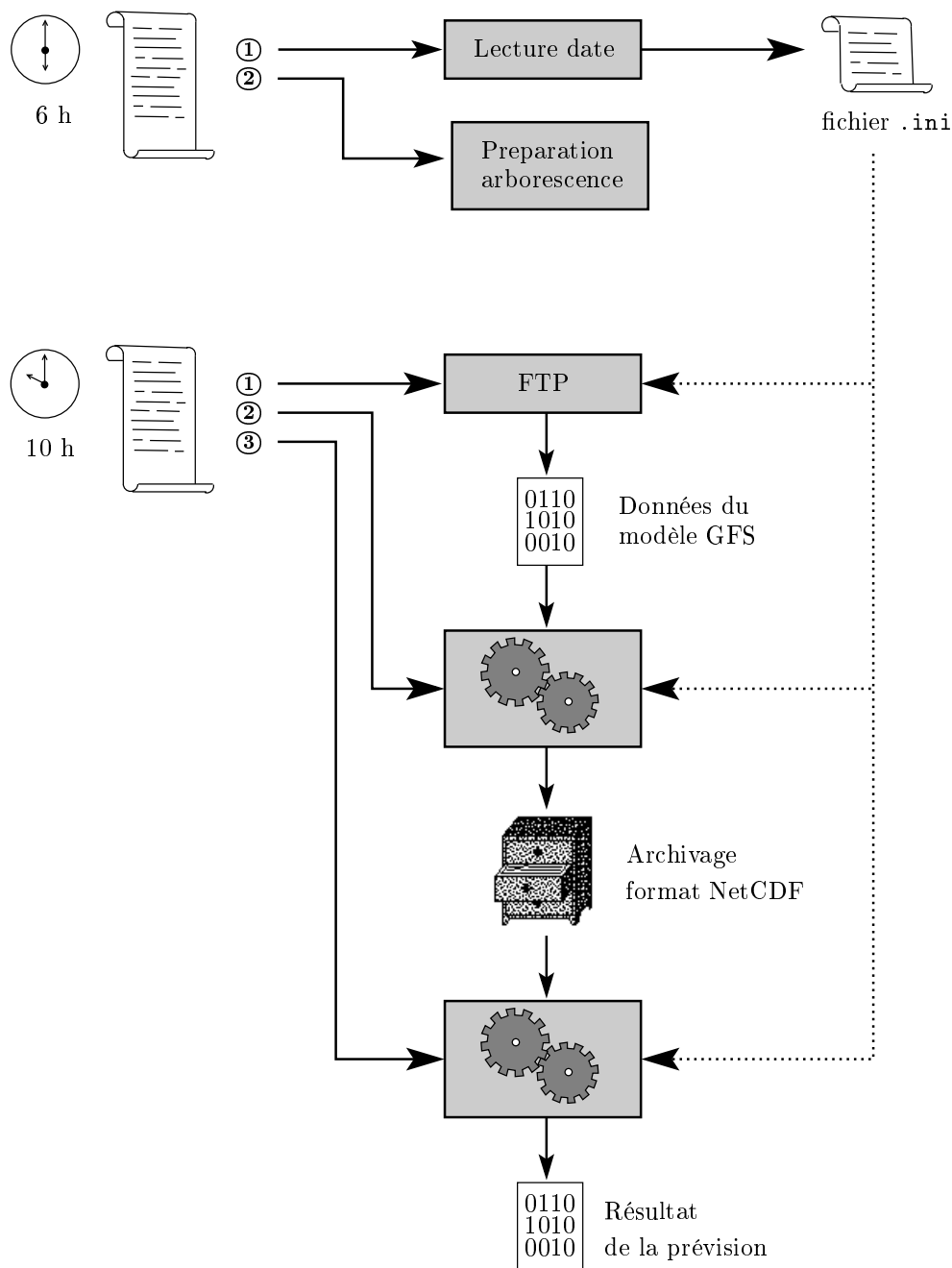


FIG. 9.1: Schéma de principe du dispositif opérationnel de prévision mis en place au LTHE.

De même, l'archivage des prévisions du modèle GFS permet de tester, ultérieurement et a posteriori, de nouveaux algorithmes dans un contexte de prévision opérationnelle.

9.1.2 L'archive de prévisions numériques

9.1.2.1 Le modèle GFS

Comme nous l'avons déjà signalé, les prévisions que nous utilisons proviennent du modèle GFS du NCEP, anciennement connu sous le nom AVN.

GFS est un modèle spectral qui considère l'ensemble du globe terrestre. Comme tout modèle opérationnel, il subit de fréquentes modifications. Il fonctionne actuellement (depuis le 29/10/02) avec une troncature T254 ($\approx 0.5^\circ$) et 64 niveaux verticaux sigma durant les 84 premières heures de prévision. Il adopte ensuite une troncature T170 et 42 niveaux jusqu'à +180 heures, puis enfin une troncature T126 avec 28 niveaux jusqu'à +384 heures. Une description complète de l'état du modèle au 1^{er} novembre 2002 est disponible à l'adresse internet : <http://wwwt.emc.ncep.noaa.gov/gmb/moorthi/gam.html> et la mise à jour des modifications apportées sur : http://wwwt.emc.ncep.noaa.gov/gmb/STATS/html/model_changes.html.

Ce modèle, dans ses versions de la fin des années 90, est celui qui a servi de base de travail pour la réalisation des réanalyses NCEP/NCAR. Nous sommes ainsi assurés d'une certaine homogénéité vis à vis de notre archive météorologique, ce qui, comme nous l'avons vu au chapitre précédent, est un pré-requis important. Cependant, au fur à mesure des évolutions de GFS, cela devrait être de moins en moins vrai.

Utilisées à des fins de recherche ou d'enseignement, les sorties de ce modèle sont disponibles gratuitement et quasiment en temps réel, via FTP anonyme à l'adresse suivante :

`ftp://ftpprd.ncep.noaa.gov/pub/data/nccf/com/avn/prod`

Cette disponibilité a été l'élément déclencheur dans le choix du modèle GFS pour l'alimentation de notre chaîne opérationnelle prototype.

9.1.2.2 Disponibilité

Initié à la mi-juin 2002, ce dispositif opérationnel fonctionnait encore au LTHE à l'automne 2003. Malheureusement, ce ne fut pas sans quelques interruptions et notre archive est un peu lacunaire. La figure 9.2 présente ainsi les périodes de fonctionnement de la chaîne opérationnelle pour l'année 2002 : sur 196 jours de fonctionnement théorique, il y a eu 11 interruptions pour un total de 18 journées de non-fonctionnement de la chaîne.

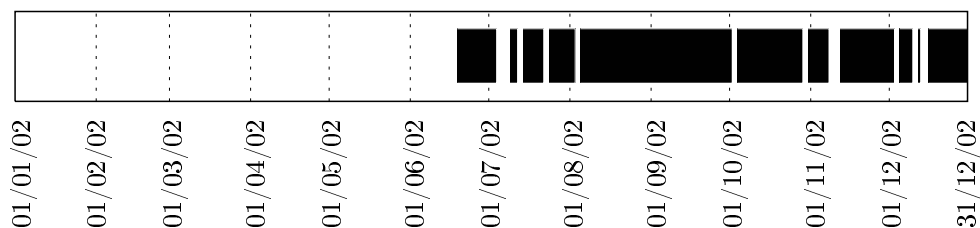


FIG. 9.2: Périodes de fonctionnement de la chaîne opérationnelle (périodes noires) pour l'année 2002.

Ces périodes de dysfonctionnement sont ainsi plus ou moins longues. Elles ont par ailleurs des causes variées, dont on n'envisageait souvent pas la possibilité, tant que l'on s'intéressait à une application théorique de la prévision :

- saturation de l'espace de stockage des données temporaires par d'autres utilisateurs,
- arrêt intempestif de l'ordinateur chargé de lancer les **scriptes de commande** par un utilisateur de passage,

- changement du système d'exploitation des serveurs,
- interruption de la connexion FTP pour cause de saturation du réseau,
- et nous avons même expérimenté le passage à l'heure d'été : la connexion FTP a alors eu lieu une heure plus tôt, avant que les données ne soient disponibles sur le serveur du NCEP, qui reste pour sa part en temps universel...

9.1.3 Performance opérationnelle des algorithmes

Parallèlement à cette archive de prévisions, nous disposons des précipitations observées sur les groupements EDF en 2002. Cette année n'avait pas été retenue dans la période de validation des algorithmes, car elle n'a pas encore été critiquée. Néanmoins, sa disponibilité nous autorise à faire une première évaluation dans un contexte opérationnel.

Nous considérons donc différentes échéances de prévision. L'échéance $J+n$ est définie comme correspondant à la prévision :

- basée sur le *run* GFS de 0hTU le jour J (donc émise à 10h15 le jour J à l'issue de notre chaîne opérationnelle),
- des précipitations qui auront lieu de 6hTU le jour $J+n$ à 6hTU le jour $J+n+1$.

Ainsi, la prévision à échéance J émise le matin du 4 septembre concerne les pluies s'étalant du 4 septembre 6hTU au 5 septembre 6hTU, alors que la prévision à échéance $J+2$ émise le même jour concerne les précipitations cumulées entre le 6 septembre 6hTU et le 7 septembre 6hTU.

Nous désirons considérer des échéances allant jusqu'à $J+6$. Or, pour que la comparaison des résultats obtenus pour différentes échéances ait un sens, il faut que les prévisions portent sur le même échantillon de situations cibles. Nous devons donc nous restreindre aux situations pour lesquelles nous disposons, sans interruption, de prévisions pour les 6 jours précédents. Cette condition réduit notre échantillon de situations cibles à seulement 116 journées sur l'année 2002.

C'est un échantillon bien court pour tirer des conclusions définitives. De plus, les données pluviométriques n'ont pas encore été critiquées. Les résultats présentés par la figure 9.3 pour deux groupements dont les données de pluie semblent de bonne qualité, n'ont donc pas d'autres prétentions que d'être considérés comme des exemples. L'extension de cette étude sur 2003 ou plus, permettra le cas échéant de généraliser les comportements observés.

On constate sur ces résultats que l'algorithme M2 présente pour l'échéance J des performances équivalentes à celles de la prévision parfaite. A $+12$ ou $+24$ heures d'échéance, les incertitudes sur les prévisions de géopotential du modèle GFS restent donc inférieures au seuil de discrimination de la recherche d'analogie : M2 trouve les mêmes analogues, que la situation cible soit prévue ou analysée. Ensuite, avec l'augmentation de l'échéance de prévision, la qualité des prévisions numériques diminue, ce qui entraîne une décroissance relativement linéaire des performances de M2 jusqu'à $J+4$. Au delà, la baisse de performance devient plus rapide, pour finalement présenter sur notre échantillon des compétences négatives vers $J+5$ ou $J+6$.

Il est important de noter que, quelle que soit l'échéance, la qualité intrinsèque de la prévision issue de l'adaptation est la même vis à vis de la situation prévue par le modèle numérique

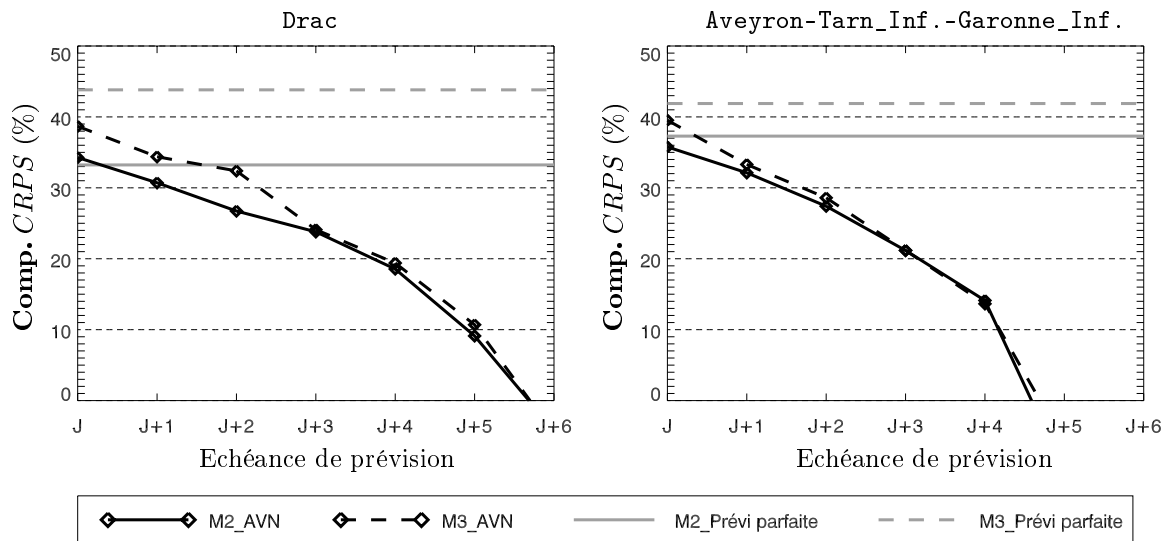


FIG. 9.3: Evolution des performances obtenues par les algorithmes M2 et M3 en fonction de l'échéance de prévision, dans le cadre d'une utilisation opérationnelle (basée sur le modèle AVN) sur 116 journées de l'année 2002 (courbes noires). Les courbes grisées correspondent aux performances obtenues en prévision parfaite (à partir des réanalyses NCEP/NCAR) sur le même échantillon.

(du moins en espérance). La diminution des performances est donc plutôt due au fait que ces situations prévues ne vont pas forcément correspondre à la situation réelle le jour J, et ceci de façon d'autant plus marquée que l'échéance de prévision augmente.

En ce qui concerne l'algorithme M3, il apparaît que les performances obtenues pour l'échéance J sont inférieures à celles de la prévision parfaite. Ainsi, contrairement aux prévisions de géopotentiels, les prévisions relatives à l'humidité (PWA et RHU) issues du modèle GFS présentent déjà, à +12 et +24 heures d'échéance, des incertitudes supérieures à celles des réanalyses. De plus, ces incertitudes augmentent plus rapidement que celles des géopotentiels, si bien que dès l'échéance J+3, la seconde sélection (M3) n'apporte aucune information supplémentaire par rapport à la première sélection (M2). M3 est même susceptible, dans ces conditions, de dégrader la qualité de la prévision en effectuant un ré-échantillonnage biaisé des situations sélectionnées par la première analogie.

Si l'écart entre M2 et M3 est déjà faible en prévision parfaite, l'utilisation opérationnelle de M3 peut même ne jamais être supérieure à celle de M2.

Il est donc préférable d'utiliser ces deux algorithmes en parallèle :

- M2 fournit une information robuste, mais manquant malheureusement de finesse.
- M3 peut améliorer M2 pour les premières échéances (J, J+1 et J+2), mais au risque de présenter une sensibilité accrue aux incertitudes de la modélisation numérique.

Rappelons enfin que notre échantillon de situations cibles (ici 116 journées sur 2002) est insuffisant pour généraliser ces conclusions. En outre, celles-ci sont propres au modèle GFS. Il serait très intéressant de pouvoir étendre ce type d'étude à d'autres modèles numériques, sur des périodes plus longues. La principale difficulté demeure malheureusement le libre accès aux prévisions opérationnelles de nombre de ces modèles météorologiques...

9.2 Etude de cas : l'épisode du 8 et 9 septembre 2002 sur les Cévennes

Alors que la chaîne de prévision opérationnelle du LTHE fonctionnait, un événement pluviométrique exceptionnel est survenu les 8 et 9 septembre 2002 dans le département du Gard et la région des Cévennes. Cet événement fut à l'origine de crues très importantes qui firent 24 victimes et causèrent des dommages estimés à 1.2 milliard d'euros (Huet *et al.* 2003).

Cette catastrophe va nous servir de cas d'étude pour étudier quelles sont les informations qu'une adaptation par analogie peut apporter pour aider à la prévision de ce type de cas extrême, dans un contexte opérationnel. Bien sûr, les conclusions que nous tirerons ne seront valables que pour ce cas particulier et n'ont absolument pas vocation à être généralisées en l'état.

9.2.1 Description de l'épisode

9.2.1.1 Contexte météorologique

La description du contexte météorologique de l'évènement des 8 et 9 septembre 2002 est tirée en grande partie de l'article de Delrieu *et al.* (2003), auquel le lecteur est invité à se reporter.

Ces précipitations exceptionnelles ont été provoquées par un Système Convectif de Meso-échelle (SCM) quasi-stationnaire, positionné au dessus du Gard. Engendré par des cellules convectives apparues en Méditerranée et qui ont progressé vers le Nord, ce SCM s'est formé vers 8hTU le 8 septembre. Ensuite, constamment ré-alimenté par de nouvelles cellules convectives provenant de la Méditerranée, il est resté en place jusqu'au 9 septembre vers midi.

La figure 9.4 présente, à partir des réanalyses NCEP/NCAR, les conditions synoptiques associées à cet événement. Le 8 septembre à 12hTU, une goutte froide d'altitude centrée sur l'Irlande s'étend en direction de l'Espagne, entraînant ainsi un flux de sud-ouest qui présente une diffluence marquée sur le sud-est de la France. Au niveau de la surface, un front froid ondule près des cotes atlantiques française, précédant le passage d'une dépression. La convection est apparue en avant de ce front, dans le Golfe du Lion, dans une zone caractérisée par un flux de basse couche de sud-est d'air chaud et humide. Les radiosondages effectués à Nîmes le 8 septembre ont montré que l'air au dessus des Cévennes était déjà conditionnellement instable et présentait de fortes valeurs d'eau précipitable.

Dans l'après-midi du 8 septembre, le flux de basse couche sur le Golfe du Lion s'est accéléré, augmentant la convergence de basse couche dans la région du Gard, tandis que la goutte froide d'altitude s'inclinait vers le sud-est, entraînant sur cette région un flux d'altitude provenant du sud-sud-ouest. L'augmentation de la diffluence de ce flux d'altitude a accéléré les mouvements verticaux au dessus de la zone de confluence des flux de basse couche, renforçant ainsi l'intensité des précipitations.

Le passage du front froid au dessus du Gard a partiellement contribué aux fortes précipitations, mais a aussi permis de briser la quasi-stationnarité du SCM en l'évacuant hors de la zone affectée.

9.2.1.2 Les précipitations engendrées

L'évènement pluvieux a duré environ 28 heures, de 0800 hTU le 8 septembre à 1200 hTU le 9 septembre. Il a généré des cumuls très importants, avec un maximum dépassant 600 mm entre Anduze et Alès. La figure 9.5 présente les cumuls obtenus sur l'ensemble de l'épisode par imagerie radar (après calibration à l'aide de postes pluviométriques).

L'un de nos groupements, **Gard-Céze**, bien que légèrement excentré de l'épicentre, se situe dans la zone des précipitations intenses. Nous allons donc étudier les résultats de l'application des algorithmes d'adaptation par analogie à ce groupement.

Cependant, comme nous l'avons mis en évidence dans le chapitre 7, la composition de ce groupement a été modifiée entre la période de calibration et celle de validation. L'observation de 57 mm donnée par ce groupement pour la journée du 8 septembre n'est donc pas fiable. Il semblerait que cette valeur ne représente la mesure que d'un seul poste sur les 5 qui constituent normalement ce groupement.

L'emplacement des postes constitutifs du groupement **Gard-Céze** est reporté sur la carte de la figure 9.5. Ceci nous permet d'estimer le cumul total sur l'épisode de ce groupement à environ 200 mm. Sur cette partie des Cévennes, la presque totalité des précipitations a eu lieu entre le 8 septembre à 6 hTU et le lendemain à 8 hTU, c'est à dire qu'elle est presque entièrement englobée par le cumul de la journée du 8 septembre, au sens de notre archive pluviométrique.

9.2.2 Prévisions émises le matin du 8 septembre

Nous allons étudier dans un premier temps les prévisions qui auraient pu être émises le matin du 8 septembre, avant que l'épisode ne commence, par différents algorithmes d'adaptation sur le groupement **Gard-Céze**. Ces prévisions sont basées sur le run du 8 septembre 2002 à 0 hTU du modèle GFS. Nous nous plaçons donc ici à l'échéance J.

Algorithme M0

Les prévisions issues de l'algorithme opérationnel M0 sont présentées par la figure 9.6. Même si elles reconnaissent le caractère pluvieux du 8 septembre, elles sous-estiment malheureusement fortement l'évènement. Elles amènent à se préparer pour un cumul d'environ 50 mm, ce qui est important, mais n'a rien d'exceptionnel sur ce groupement qui connaît une pluie journalière décennale de 154 mm.

Si l'on considère le détail de la distribution des précipitations prévues, on constate néanmoins que sur les 50 analogues, 2 ont engendré des précipitations quasi-décennales (de l'ordre de 140 mm). En outre, l'observation des 10 meilleures analogues semble indiquer qu'il y a deux populations dans cette distribution : une première centrée sur des précipitations d'une dizaine de millimètres, et une seconde qui regroupe plutôt des précipitations autour de 60 mm.

L'analogie reconnaît donc bien qu'il y a un potentiel de fortes précipitations, mais son manque de finesse ne permet pas d'aller plus loin et rien n'indique que la décennale pourra être dépassée.

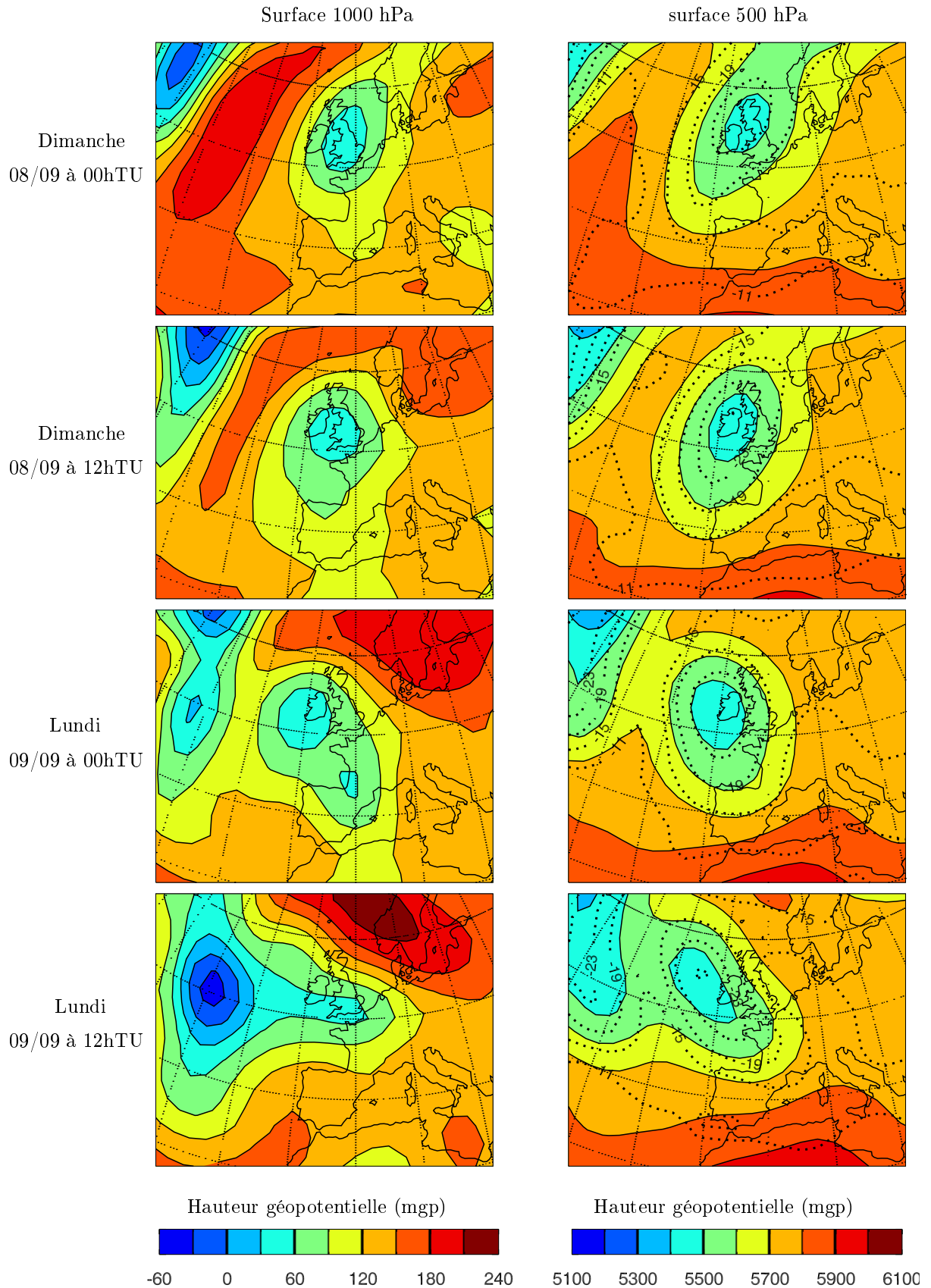


FIG. 9.4: Réanalyses NCEP/NCAR du dimanche 8 septembre 2002 à 0hTU au lundi 9 à 12hTU, par pas de 12 heures. **Panneau de gauche** : carte de géopotentiels au niveau 1000 hPa ; **panneau de droite** : carte de géopotentiels (contours pleins) et de température (contours pointillés, en °C) au niveau 500 hPa.

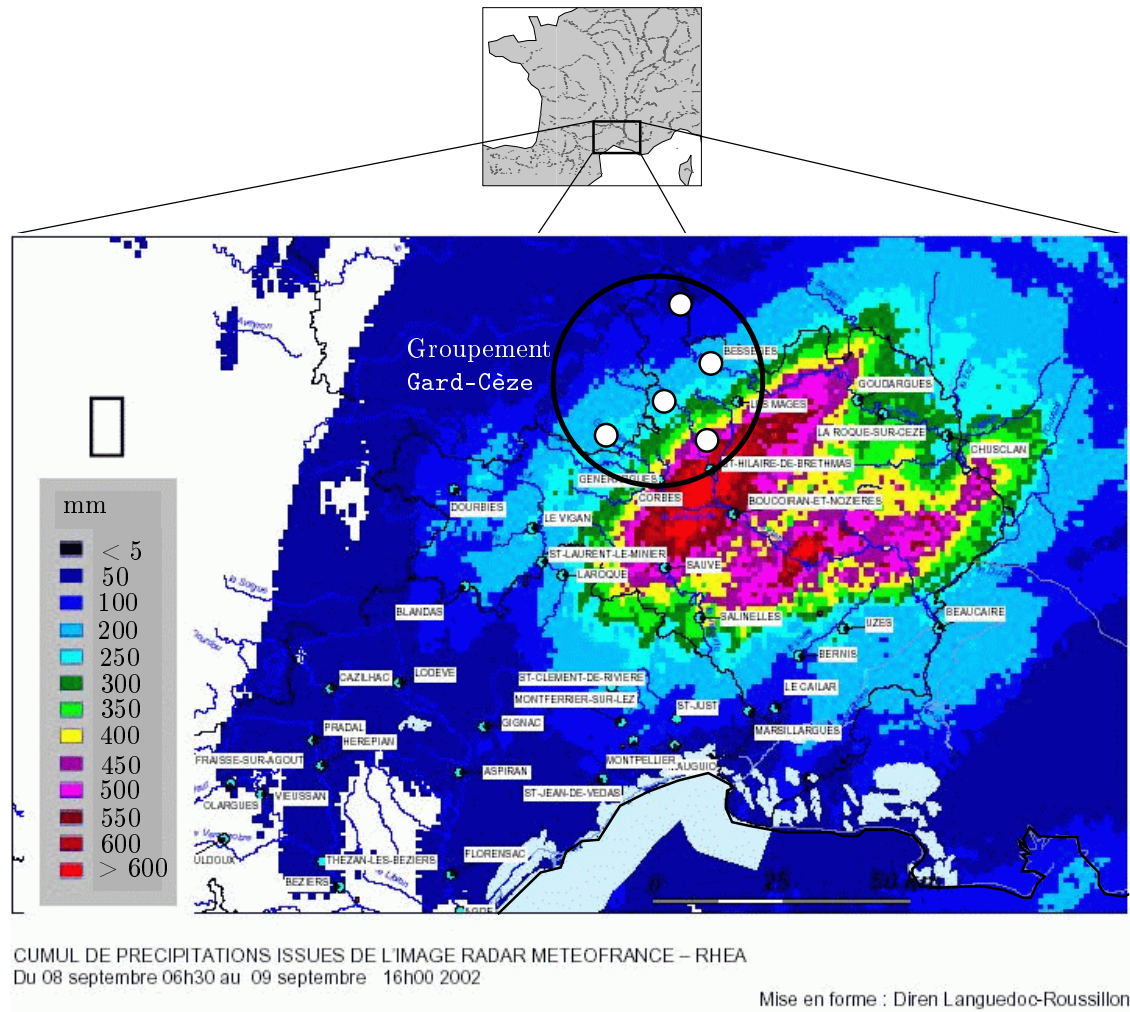


FIG. 9.5: Cumul de précipitation de l'épisode des 8 et 9 septembre 2002 (mesures radars calibrées, source : Diren Languedoc-Roussillon).

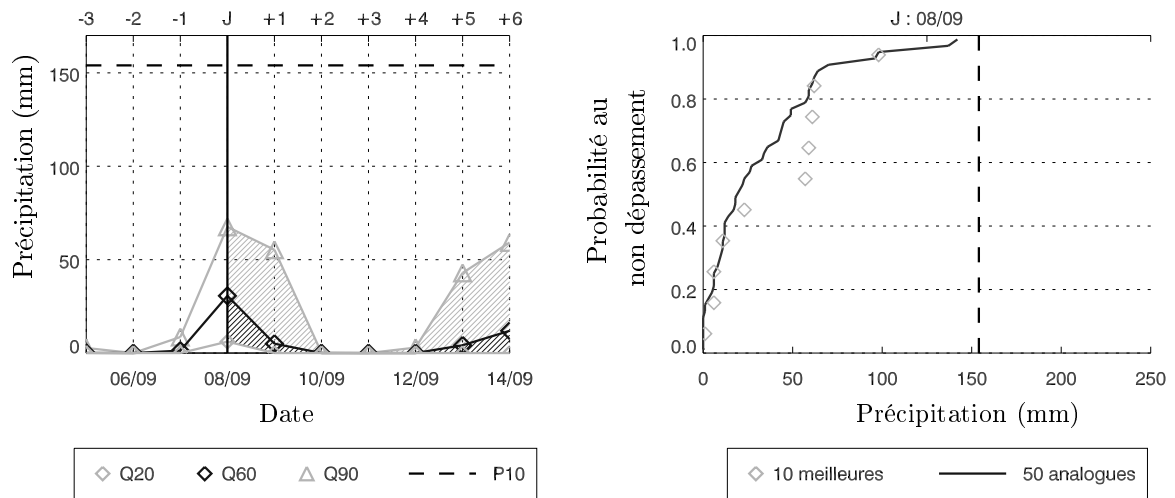


FIG. 9.6: Prévisions émises le matin du 8 septembre 2002 par l'algorithme M0 concernant le groupement Gard-Cèze. **Panneau de gauche** : évolution des quantiles $Q_{20\%}$, $Q_{60\%}$ et $Q_{90\%}$ en fonction de l'échéance de prévision (on rappelle les prévisions émises les trois jours précédents pour l'échéance J) ; **panneau de droite** : distribution des précipitations prévues pour la journée du 8 septembre (échéance J).

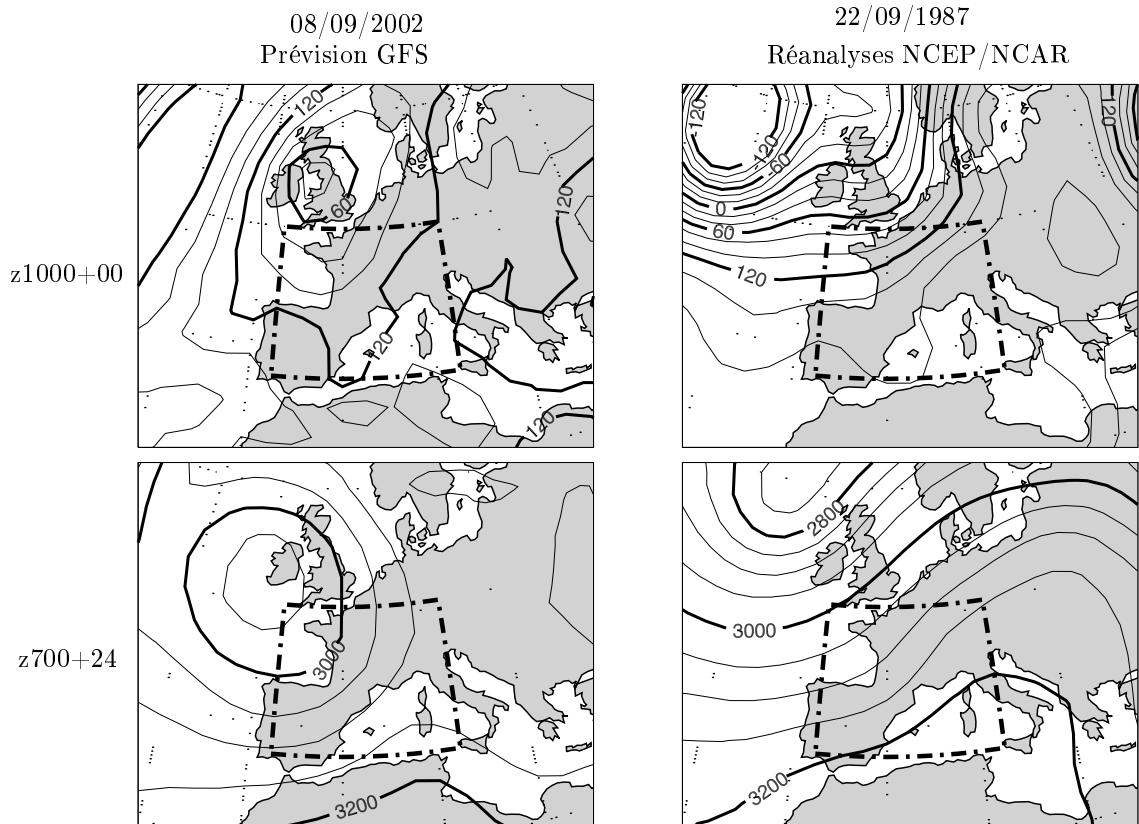


FIG. 9.7: Comparaison des géopotentiels $z_{1000+00}$ et z_{700+24} , entre la situation cible prévue par GFS pour le 08/09/02 (échéance J) et la meilleure analogue au sens de M0 : le 22/09/87. Le cadre tireté représente la fenêtre d'analogie utilisée.

La figure 9.7 présente la comparaison des géopotentiels $z1000+00$ et $z700+24$ de la situation du 8 septembre telle que prévue par le modèle AVN (situation cible), avec ceux de la situation du 22 septembre 1987, reconnu comme étant la meilleure analogue au sens de M0. Les précipitations associées à cette situation sont d'à peine 6.6 mm sur le groupement **Gard-Cèze**.

Les champs $z1000+00$ et $z700+24$ font partie de ceux sur lesquels se base l'analogie dans le cas de l'algorithme M0. On constate que, au moins à l'intérieur de la fenêtre d'analogie, la similitude de forme est tout à fait satisfaisante, surtout pour $z1000+00$. Par contre, si l'on considère le champ $z1000+12$ (figure 9.8), qui s'est révélé être l'un des champs les plus informatifs (cf. chapitre 4), on ne retrouve pas le même niveau de similitude. L'analogie de M0 se « *disperse* » entre quatre champs, fournissant ainsi des analogues globalement correctes, mais qui ne présentent pas pour autant une forte analogie sur les champs les plus pertinents. Il est en outre frappant de voir que, à plus grande échelle, les centres d'actions de cette meilleure analogue au sens de M0 (zones de hautes et basses pression, thalweg, ...) ne correspondent absolument pas à ceux de la situation cible.

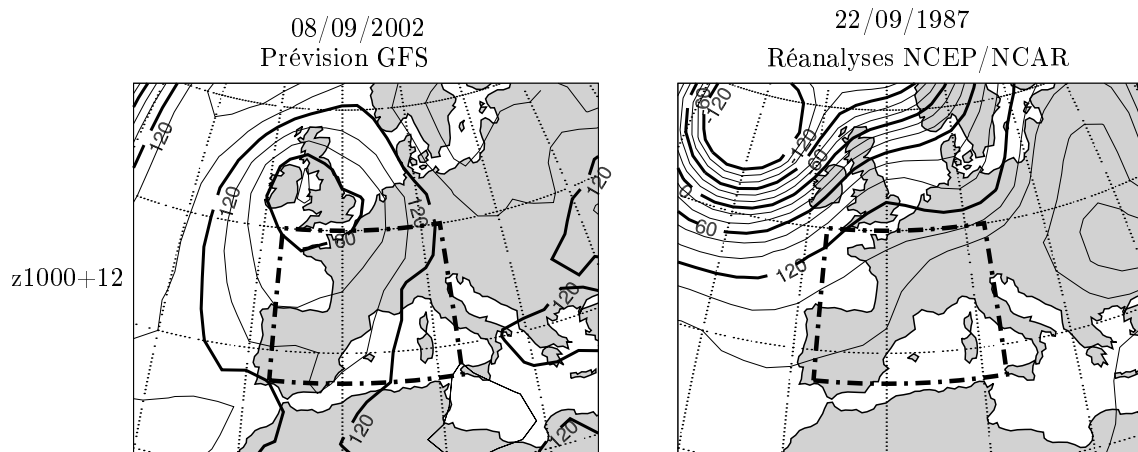


FIG. 9.8: Comparaison des géopotentiels $z1000+12$, entre la situation cible prévue par GFS pour le 08/09/02 (échéance J) et la meilleure analogue au sens de M0 : le 22/09/87. Le cadre tireté représente la fenêtre d'analogie utilisée.

M2bis : Utilisation de $z1000+12$ & $z500+24$

Nous utilisons maintenant un algorithme analogue à M2, en employant la combinaison de géopotential $z1000+12$ & $z500+24$. La différence est due au fait que la fenêtre d'analogie a été optimisée spécifiquement pour le groupement **Gard-Cèze** sur la période de calibration, ce qui nous a conduit à considérer une fenêtre légèrement plus étendue que celle de M0. Nous appelons cet algorithme M2bis.

Les prévisions de M2bis, présentées par la figure 9.9, font état de précipitations un peu plus soutenues que M0. Elles incitent plutôt à faire face à des cumuls de l'ordre de 60 mm. Elles montrent par ailleurs de manière plus claire l'arrêt des précipitations pour le 9 septembre.

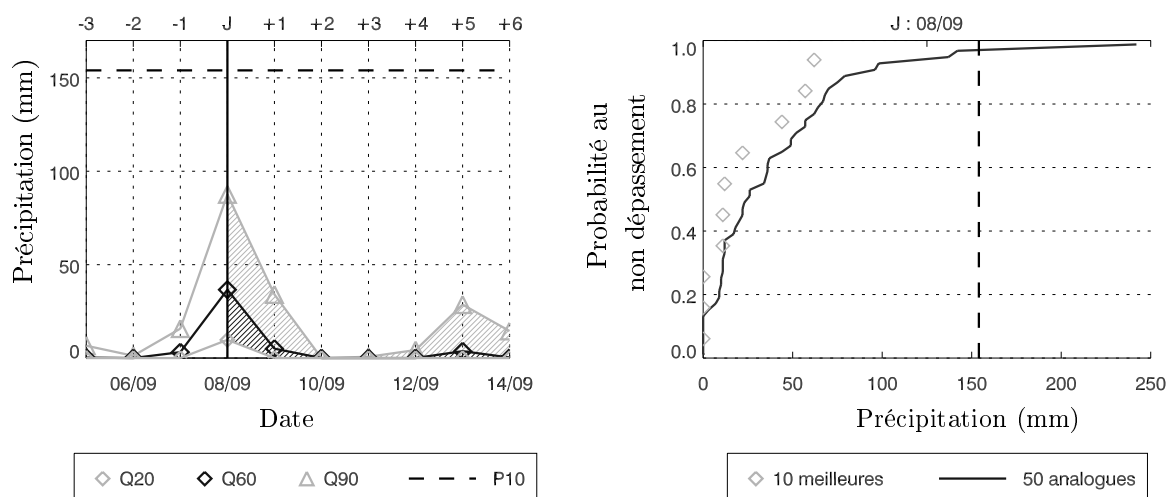


FIG. 9.9: Prévisions émises le matin du 8 septembre 2002 par l'algorithme M2bis concernant le groupement Gard-Cèze. **Panneau de gauche** : évolution des quantiles $Q_{20\%}$, $Q_{60\%}$ et $Q_{90\%}$ en fonction de l'échéance de prévision (on rappelle les prévisions émises les trois jours précédents pour l'échéance J) ; **panneau de droite** : distribution des précipitations prévues pour la journée du 8 septembre (échéance J).

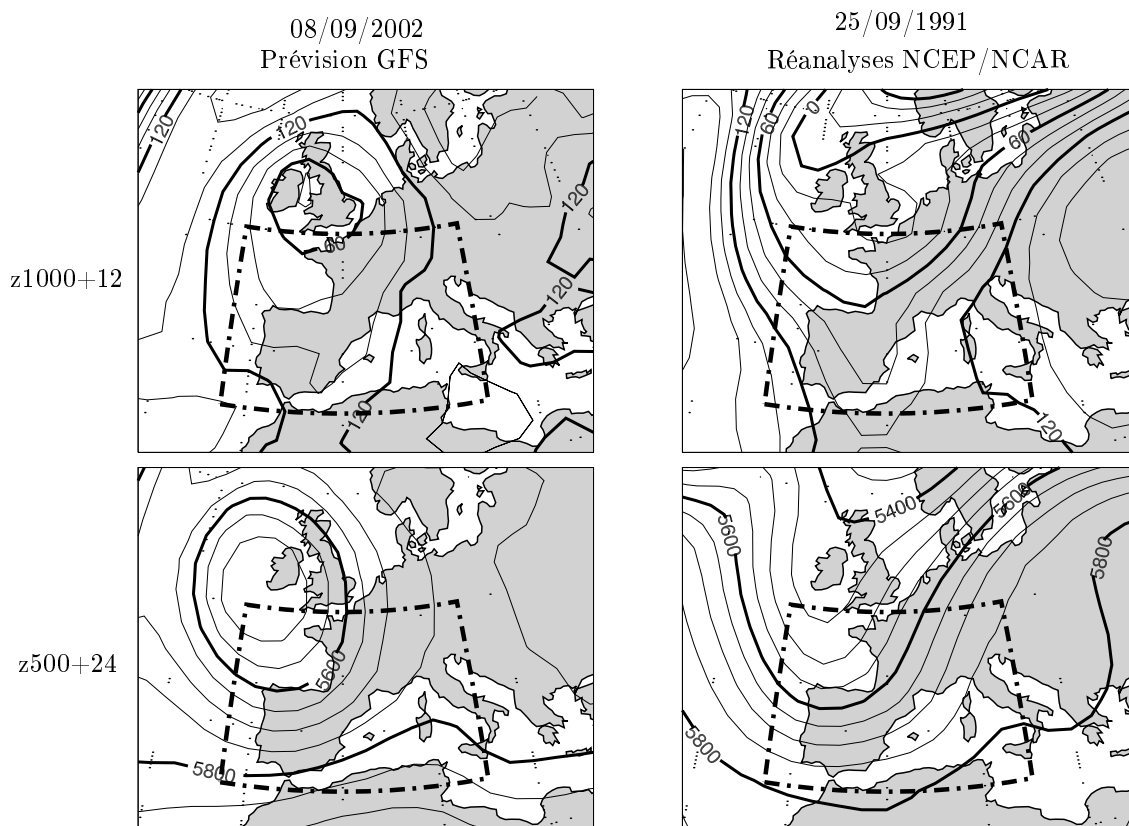


FIG. 9.10: Comparaison des géopotentiels $z_{1000+12}$ et z_{500+24} , entre la situation cible prévue par GFS pour le 08/09/02 (échéance J) et la meilleure analogue au sens de M2bis : le 25/09/91. Le cadre tireté représente la fenêtre d'analogie utilisée.

La distribution des précipitations prévues pour l'échéance J reste toutefois assez comparable à celle de M0, et la distribution des 10 meilleures analogues semble encore montrer la présence de deux populations distinctes.

La situation du 25 septembre 1991, qui a engendré 57 mm de précipitation, est celle qui a été reconnue comme la meilleure analogue par M2bis (figure 9.10). On constate que la similitude est en effet très forte à l'intérieure de la fenêtre d'analogie pour $z1000+12$ et $z500+24$, qui sont maintenant les variables d'analogie. Néanmoins, les centres d'action de cette situation sont encore notablement différents de ceux de la situation cible.

Cependant, on remarque que parmi les 50 analogues retenues, on trouve à la 47^{ème} place la situation qui a engendré les plus fortes précipitations de notre archive sur le groupement Gard-Céze : le 30 septembre 1958 avec 242.4 mm. Cette situation, qui fut à l'origine de terribles crues, présente effectivement une analogie de moins bonne qualité que le 25/09/91 avec la situation cible, mais montre par contre une plus forte similitude dans la disposition des centres d'actions (figure 9.11). Elle est donc finalement peut-être plus voisine du 8 septembre 2002 que ne le laisse entendre sa 47^{ème} place...

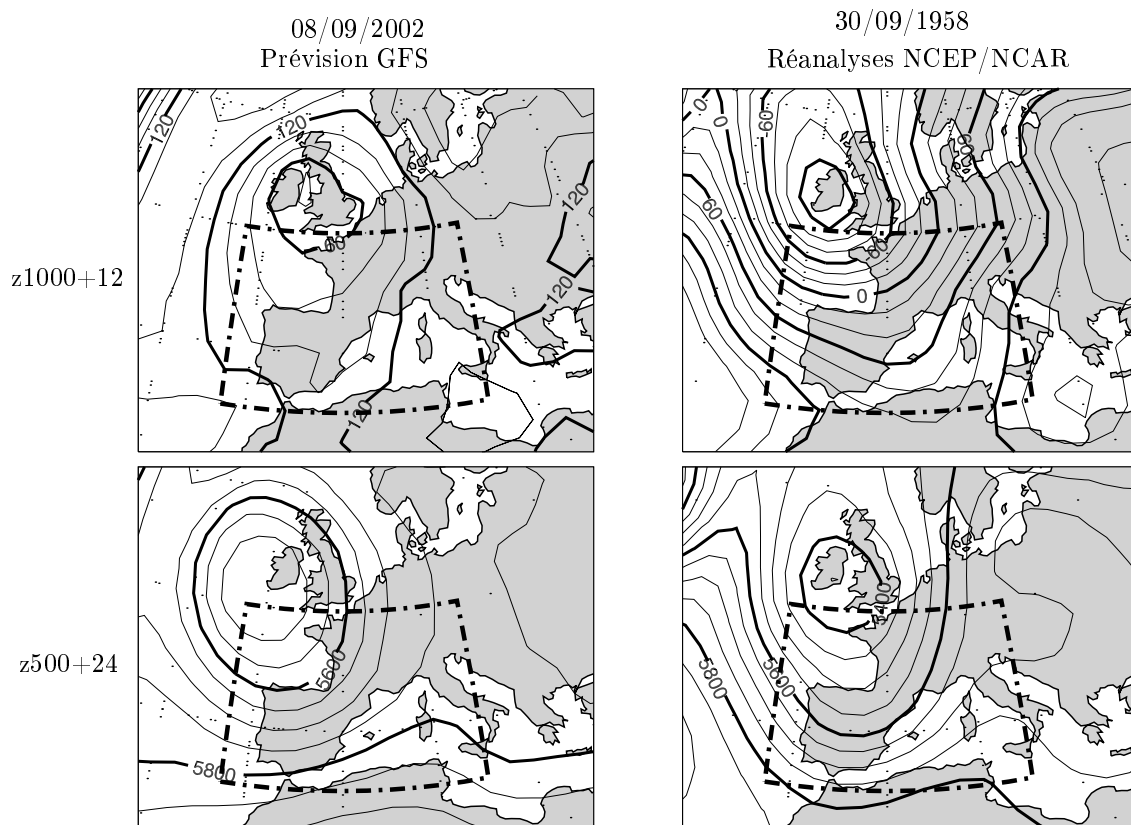


FIG. 9.11: Comparaison des géopotentiels $z1000+12$ et $z500+24$, entre la situation cible prévue par GFS pour le 08/09/02 (échéance J) et le 30/09/58 (plus fortes précipitations de l'archive). Le cadre tireté représente la fenêtre d'analogie utilisée.

L'utilisation de M2bis, malgré l'apparition de la situation du 30 septembre 1958 dans le panel des situations analogues, ne permet pas encore d'anticiper l'événement exceptionnel qui

se prépare, à cause de son manque de finesse. De même que M0, M2bis n'arrive pas à faire de discrimination à l'intérieur des deux populations présentes au sein des analogues.

M2_dyn : Fenêtre d'analogie dynamique

Nous avons par ailleurs testé l'utilisation de fenêtres dynamiques sur ce cas d'étude, par l'intermédiaire de l'algorithme M2_dyn. Basé sur M2, il considère des fenêtres dynamiques telles que décrites dans le chapitre 5 (cf. 5.2.2), en utilisant des paramètres optimisés sur la période de calibration : LAMONT = 4, LAVAL = 1, THP = 1 et TBP = 0.

Comme nous l'avons déjà vu, ces algorithmes ne permettent pas d'obtenir les performances optimales dans le cas général. La figure 9.12 montre néanmoins que dans le cas précis du 8 septembre 2002, les prévisions obtenues sont supérieures à celles issues des algorithmes statiques (M0 et M2bis). Le quantile $Q_{60\%}$ atteint 50 mm et le quantile $Q_{90\%}$ dépasse les 100 mm. Or, dans ces conditions, le quantile de référence est plutôt $Q_{90\%}$ (cf. chapitre 7). A la vue de cette prévision, on s'attend donc à observer des cumuls de l'ordre de 100 mm.

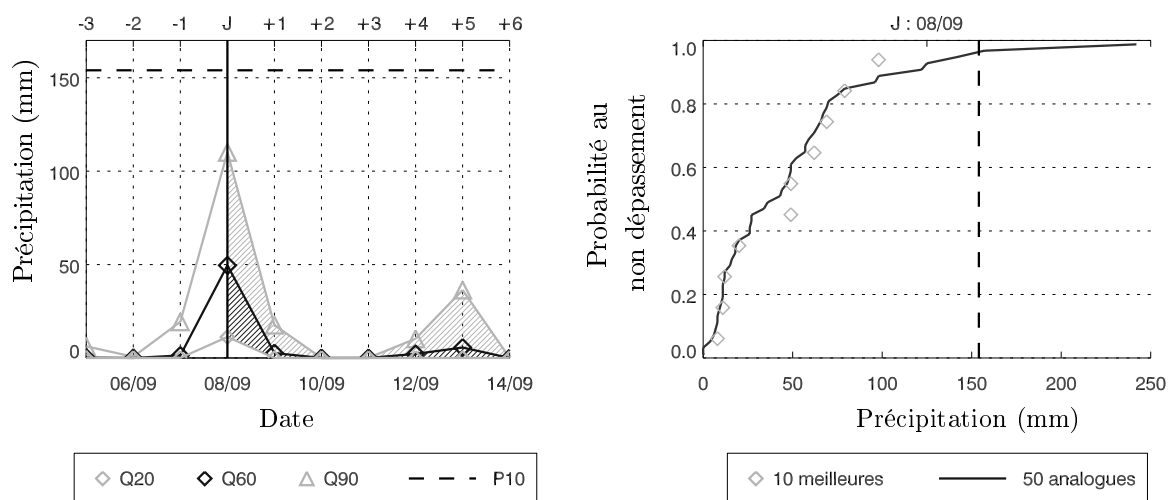


FIG. 9.12: Prévisions émises le matin du 8 septembre 2002 par l'algorithme M2_dyn concernant le groupement Gard-Cèze. **Panneau de gauche :** évolution des quantiles $Q_{20\%}$, $Q_{60\%}$ et $Q_{90\%}$ en fonction de l'échéance de prévision (on rappelle les prévisions émises les trois jours précédents pour l'échéance J) ; **panneau de droite :** distribution des précipitations prévues pour la journée du 8 septembre (échéance J).

Ainsi que le montre la distribution des précipitations prévues pour l'échéance J, M2_dyn donne un poids plus important que M0 ou M2bis à la population des fortes précipitations : parmi les 10 meilleures analogues, 6 dépassent les 50 mm. En outre, à l'intérieur de l'ensemble de la distribution, on recense 2 situations qui donnèrent lieu à des précipitations plus que décennales, dont à nouveau le 30 septembre 1958.

M2_dyn annonce donc l'occurrence de très fortes précipitations, de l'ordre de 100 mm, avec un risque de dépassement de ces valeurs. Cependant, cet algorithme ne permet pas d'annoncer à coup sûr un événement exceptionnel. De plus, les performances des fenêtres dynamiques dans le

cas général n'incitent pas vraiment, a priori, à donner la préférence à cette prévision plutôt qu'à celle de M0 ou de M2bis.

On constate sur la figure 9.13 que la fenêtre d'analogie dynamique englobe d'elle même la diffuence du flux d'altitude, qui est l'une des caractéristiques importantes de la situation synoptique du 8 septembre 2002. La meilleure analogue au sens de M2_dyn est le 18 août 1958, qui a engendré 12.4 mm de pluie sur le groupement Gard-Céze.

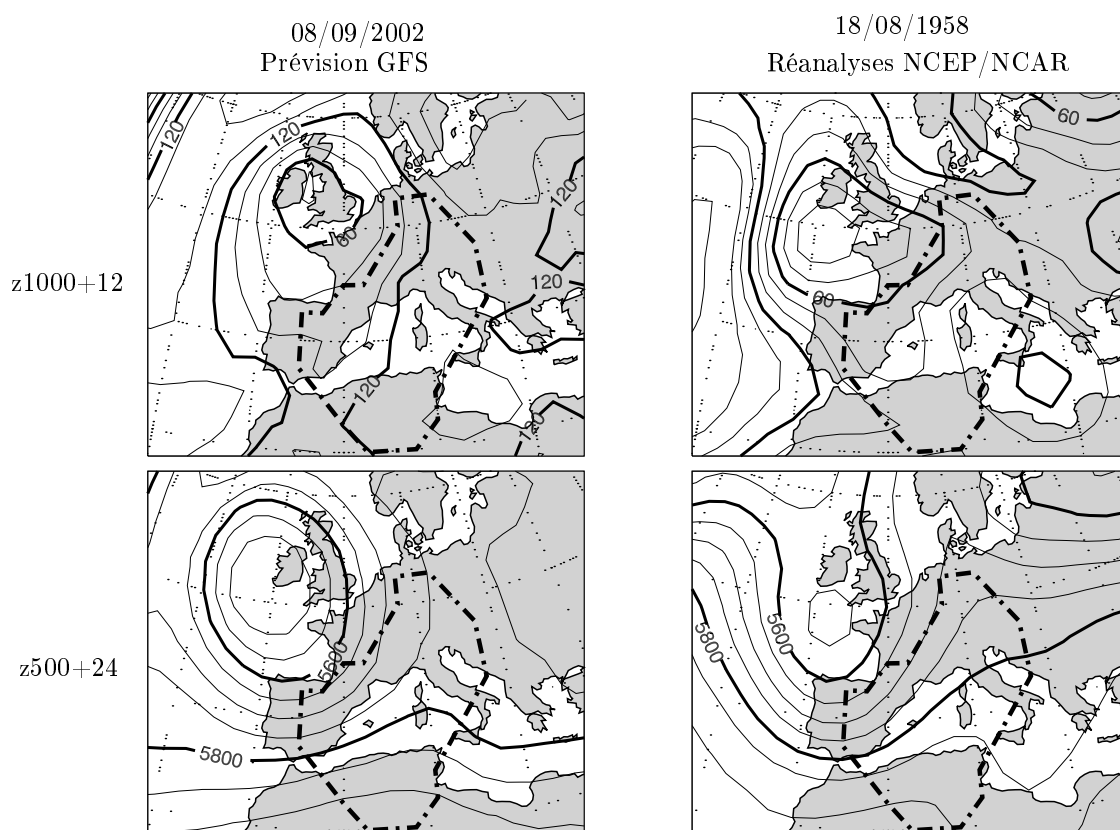


FIG. 9.13: Comparaison des géopotentiels $z1000+12$ et $z500+24$, entre la situation cible prévue par GFS pour le 08/09/02 (échéance J) et la meilleure analogue au sens de M2_dyn : le 18/08/58. Le cadre tireté représente la fenêtre d'analogie utilisée.

M3bis : Ajout d'une analogie de second ordre

Enfin, nous avons utilisé un algorithme M3bis qui effectue deux recherches successives d'analogie. La première sélection porte sur la circulation générale, de façon identique à M2bis, et la seconde sur les variables d'humidités locales, tel que l'effectue M3.

Les prévisions issues de M3bis pour la journée du 8 septembre 2002 (à échéance J), sont bien plus fortes que toutes celles que nous avons considérées jusqu'ici (figure 9.14). Le quantile $Q_{20\%}$ dépasse les 20 mm, le $Q_{60\%}$ atteint 60 mm et surtout, $Q_{90\%}$ est quasiment au niveau de la pluie décennale. De telles valeurs sont très rares. En effet, dans une archive de 44 années, on dispose en moyenne de seulement 4 ou 5 observations qui dépassent la valeur décennale. Il n'est donc pas possible d'en retenir un plus grand nombre, ce qui dans une distribution de 30 situations, limite fortement les valeurs que peuvent atteindre les différents quantiles.

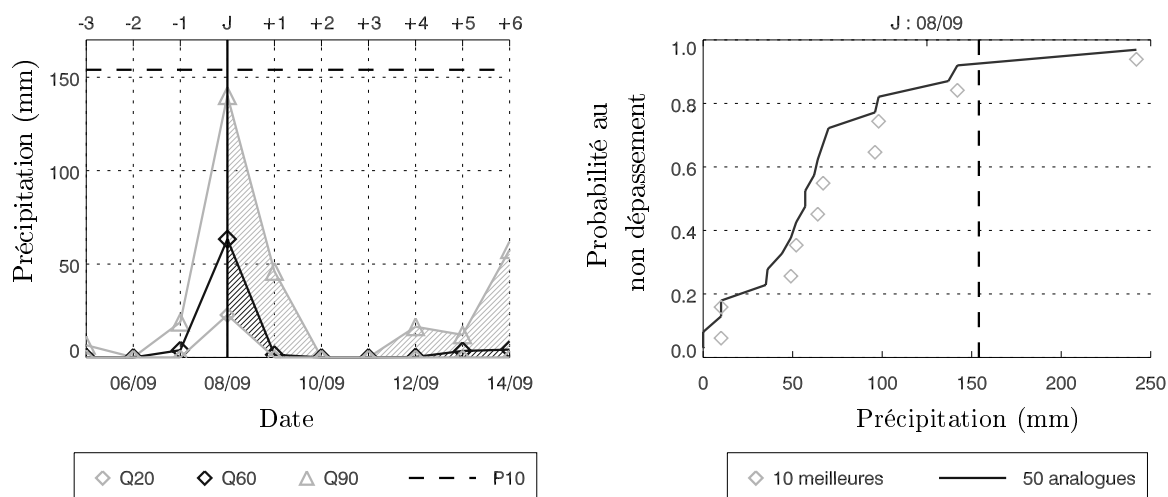


FIG. 9.14: Prévisions émises le matin du 8 septembre 2002 par l'algorithme M3bis concernant le groupement Gard-Cèze. **Panneau de gauche** : évolution des quantiles $Q_{20\%}$, $Q_{60\%}$ et $Q_{90\%}$ en fonction de l'échéance de prévision (on rappelle les prévisions émises les trois jours précédents pour l'échéance J) ; **panneau de droite** : distribution des précipitations prévues pour la journée du 8 septembre (échéance J).

Cette hausse des quantiles n'est pas due au fait que l'on passe de 50 analogues avec M2bis à seulement 30 avec M3bis. Si l'on ne retient effectivement que les 30 meilleures analogues de M2bis, $Q_{90\%}$ ne dépasse alors pas 80 mm. La hausse des quantités de pluie prévues par M3bis est donc bien le fait de l'information contenue dans les variables d'humidités, lesquelles connaissent de fortes valeurs, très éloignées de la moyenne, pour la situation du 8 septembre 2002.

On sait aussi que lorsque d'importantes précipitations sont prévues, il est préférable de s'intéresser au $Q_{90\%}$ plutôt qu'au $Q_{60\%}$. Cependant, pour de tels seuils de précipitations, même le quantile $Q_{90\%}$ possède une très faible probabilité de détection (POD). Il est donc fort probable que les précipitations réelles soient encore plus importantes que $Q_{90\%}$, et dépassent donc la décennale.

La distribution des précipitations prévues montre que parmi les deux populations retenues par M2bis, M3bis ne garde que les situations appartenant à la population qui engendre de fortes précipitations : 8 des 10 meilleures analogues dépassent les 50 mm. En outre, à l'issue de la seconde sélection, le 30 septembre 1958 passe de la 47^{ème} à la 3^{ème} place. La plus forte précipitation de l'archive est maintenant reconnue comme étant la troisième meilleure analogue à la situation du 8 septembre 2002. Voilà qui permet de craindre le pire...

Les prévisions de M3bis sont donc suffisantes, si ce n'est pour déclencher une alerte, au moins pour alarmer le prévisionniste et l'inciter à approfondir la situation. Le principal inconvénient de l'adaptation pour de tels événements, est que la prévision n'est plus quantitative : on sait que les précipitations ont de fortes chances d'être extrêmes, au delà de la pluie décennale, mais les limitations de l'archive pluviométrique nous empêchent d'estimer de combien la pluie décennale

sera dépassée. Toutefois, existe-t-il des méthodes parfaites pour anticiper des événements aussi rares ?

La meilleure analogue au sens de M3bis est le 4 novembre 1955, qui a engendré une pluie de 49.7 mm (figure 9.15). L'analogie synoptique est en effet moins satisfaisante que pour le 25 septembre 1991 (meilleure analogue de M2bis) sur la fenêtre d'analogie, mais les centres d'actions sont, de même que pour le 30 septembre 1958, plus en phase avec ceux de la situation cible.

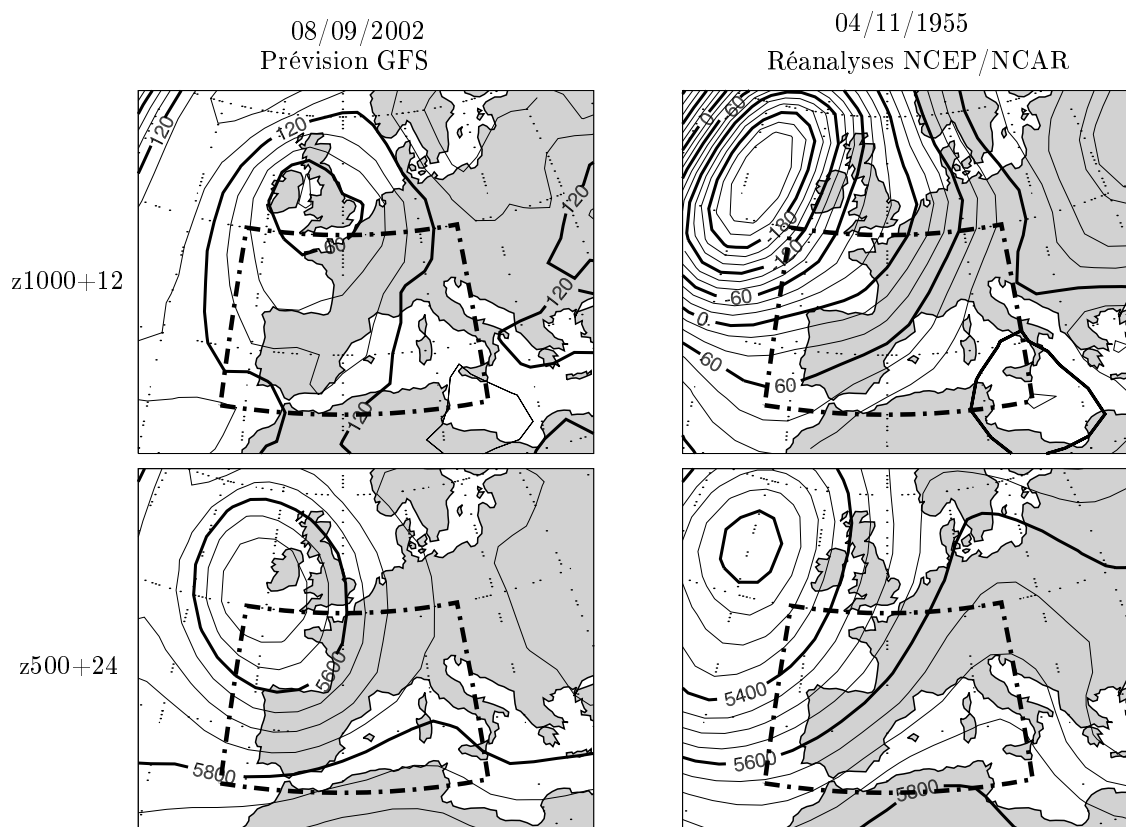


FIG. 9.15: Comparaison des géopotentiels $z1000+12$ et $z500+24$, entre la situation cible prévue par GFS pour le 08/09/02 (échéance J) et la meilleure analogue au sens de M3bis : le 04/11/55. Le cadre tireté représente la fenêtre d'analogie utilisée.

Il est à noter que de telles valeurs de quantile $Q_{90\%}$ sont très rares : si l'on se réfère aux 44 années de la période de calibration, seules 4 situations ont affiché des $Q_{90\%}$ supérieurs à 130 mm (en prévision parfaite). Cette rareté, que l'on retrouve ici, doit donc encore plus nous inciter à la prudence !

9.2.3 Evolution des prévisions dans les jours précédents

L'algorithme M3bis permet donc d'annoncer l'imminence d'un événement très important le 8 septembre au matin pour les prochaines 24 heures. Mais était-il possible d'avoir une anticipation plus importante encore ?

La figure 9.16 présente les prévisions émises par les algorithmes M0 et M3bis, du 3 au 8 septembre (de J-5 à J).

Le mardi 3 septembre, soit 5 jours avant l'évènement, les deux algorithmes prévoient une fin de semaine parfaitement sèche. Les prévisions du modèle GFS vont finalement se révéler tout à fait inexactes...

Le lendemain (mercredi 4 septembre, J-4) et le surlendemain (jeudi 5 septembre, J-3), le modèle GFS a changé d'avis et les deux algorithmes prévoient maintenant de faibles précipitations pour le 8 septembre.

A partir du vendredi 6 septembre (J-2), les précipitations prévues restent toujours faibles, mais M3bis tend à se démarquer de M0 en annonçant des précipitations plus fortes. Le modèle GFS cerne encore mal la situation synoptique du 8 septembre, mais a déjà reconnu que l'humidité sera élevée sur le sud-est de la France.

La veille de l'évènement, le samedi 7 septembre (J-1), le modèle GFS change encore d'avis. M0 prévoit maintenant des précipitations moyennes, tandis que M3bis annonce de très fortes précipitations, avec un quantile $Q_{90\%}$ de 130 mm. Ces prévisions sont peu différentes de celles émises ensuite le matin du 8 septembre : celles de M3bis sont déjà suffisantes le samedi 7 au matin pour alerter un prévisionniste !

L'adaptation du modèle GFS par M3bis permet donc d'anticiper l'épisode du 8 septembre 2002 d'au moins 24 heures. Dans ce cas particulier, la limitation des capacités d'anticipation n'est pas liée aux incertitudes qui entachent en général les prévisions d'humidité dès que les échéances s'éloignent. A notre étonnement, ce sont plutôt les prévisions portant sur la circulation générale qui ne se sont révélées satisfaisantes qu'à partir de J-1.

9.3 Conclusion

Ce chapitre nous a permis de tester l'application de nos algorithmes d'adaptation dans un contexte de prévision opérationnelle.

Pour ce faire, nous avons implémenté au LTHER une chaîne opérationnelle qui permet d'avoir accès tous les jours aux sorties du modèle GFS du NCEP. A l'issue de cette chaîne, les algorithmes d'adaptation se basent sur ces prévisions pour caractériser les 6 jours suivants et émettre des prévisions de précipitations sur l'ensemble des groupements de l'archive EDF. Un outil de visualisation graphique permet ensuite de consulter ces prévisions à tout moment.

L'archivage des prévisions nous a autorisé à faire une évaluation des performances opérationnelles de l'adaptation sur 116 journées de 2002. Il est ainsi apparu qu'à l'échéance de prévision J (c'est à dire pour les 24 heures à venir), un algorithme basé sur les seuls géopotentiels, tel que M2, présente le même niveau de performance qu'en prévision parfaite. Les algorithmes qui prennent aussi en compte les variables d'humidité, tels que M3, présentent par contre des performances inférieures à la prévision parfaite, du fait de la plus grande incertitude entachant les prévisions d'humidité. Si la différence de performance entre les algorithmes en prévision parfaite

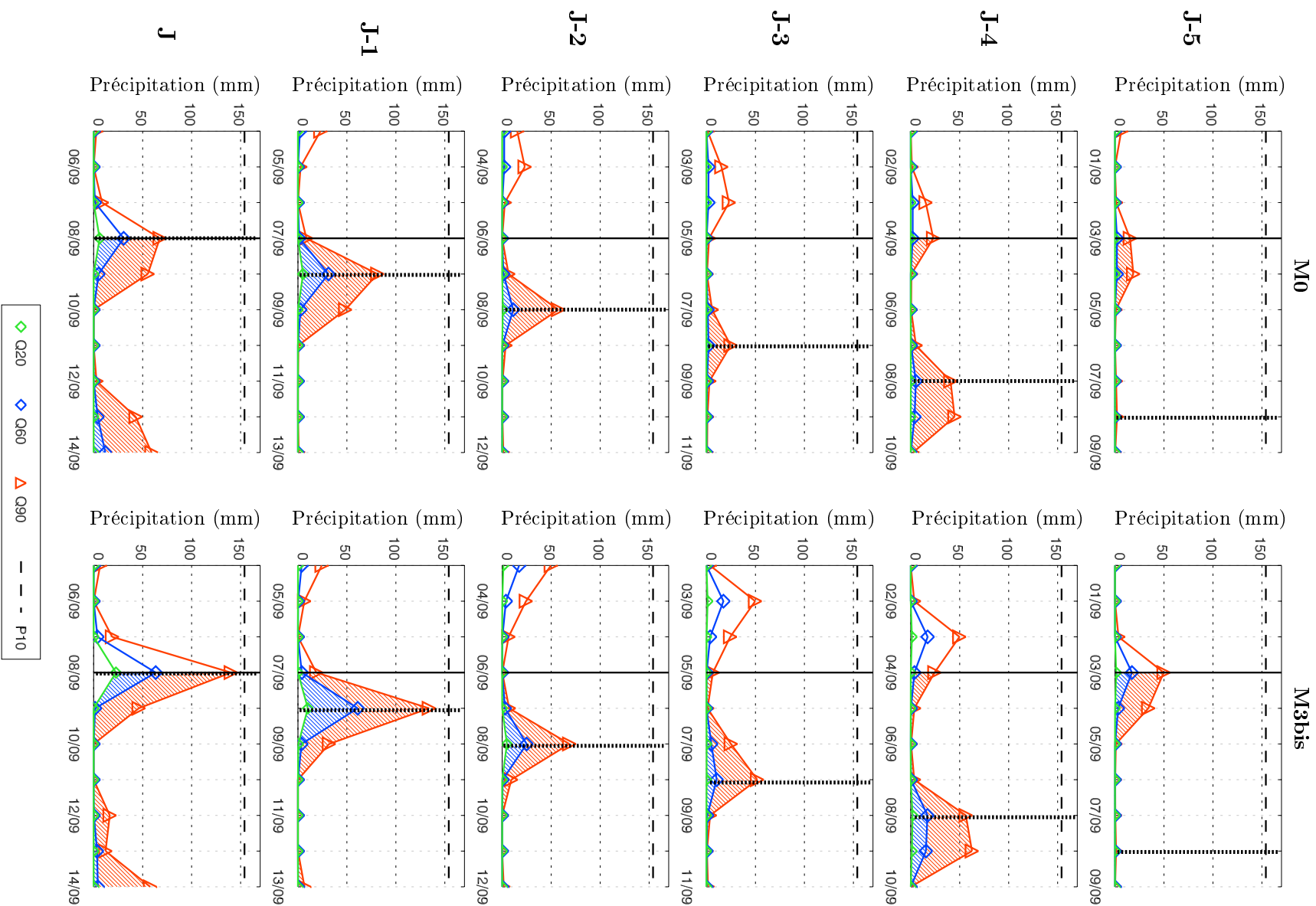


Fig. 9.16: Prévisions émises par M0 et M3bis du 3 au 8 septembre 2002 (soit de J-5 à J par rapport à l'évènement du 8 septembre 2002).

est suffisamment importante, M3 affiche des résultats supérieurs à ceux de M2 sur les premières échéances. Au delà de J+2, la seconde sélection n'apporte plus d'information supplémentaire.

Nous nous sommes ensuite intéressés à l'étude, dans un contexte opérationnel, de la prévision des précipitations intenses survenues les 8 et 9 septembre 2002 dans le département du Gard, en nous préoccupant plus particulièrement du groupement **Gard-Cèze**.

Le matin de l'évènement, avant que les précipitations ne commencent, les algorithmes basés sur les seuls géopotentiels ne prévoyaient que des précipitations moyennes. Néanmoins, il était possible de distinguer deux populations différentes au sein des analogues retenues, l'une d'elle étant susceptible d'engendrer de fortes précipitations. Un algorithme à deux sélections s'est avéré capable d'effectuer une discrimination parmi ces deux populations en ne conservant que celles qui étaient associées aux plus fortes précipitations. Ainsi, la prévision de M3bis présente un quantile $Q_{90\%}$ qui atteint quasiment la valeur de la précipitation journalière décennale. Etant donné les limitations de l'archive pluviométrique, on peut difficilement espérer obtenir plus. Cette prévision annonce donc l'évènement catastrophique qui se prépare, d'autant plus que la troisième meilleure analogue n'est autre que la situation qui a produit les plus fortes précipitations de nos 44 années d'archive.

En considérant des échéances plus éloignées, il est apparu que M3bis était capable d'annoncer l'évènement des 8 et 9 septembre dès le matin du 7 septembre (J-1). Cependant, pour les échéances plus lointaines, les prévisions sont complètement fantaisistes. Il semblerait que le modèle GFS n'ait pas réussi à prévoir les conditions synoptiques de façon satisfaisante avant cette échéance.

Cette étude de cas a permis de mettre en lumière l'un des points faibles de nos algorithmes d'analogie. En effet, le respect de la similitude sur la fenêtre d'analogie n'implique pas pour autant que la disposition des centres d'actions de la circulation générale soit comparable entre la situation cible et la situation candidate. Or, les situations conservées par l'algorithme à deux sélections, sur la base de critères d'humidité, sont justement celles qui respectent la disposition des centres d'action de la situation cible. Il pourrait donc être intéressant d'exprimer un critère d'analogie sous la forme d'une schématisation de ces centres d'action...

D'autres études de cas deviendront disponibles au fur et à mesure que les évènements se produiront. Ces retours d'expériences permettront d'enrichir notre expérience quant aux capacités de l'adaptation par analogie et peut-être d'identifier les comportements généralisables.

Conclusions

« *Enfin, quand il pleut, ça mouille.
C'est aussi simple que ça.* »

Réplique de Magnolia, film de Paul Thomas Anderson

Le travail présenté ici participe à l'étude de l'application du concept d'*analogie* à la prévision et à la simulation des précipitations. Notre approche a consisté, à partir de la caractérisation d'une situation météorologique cible pour laquelle on souhaite « *prévoir* » les précipitations associées, à rechercher dans une archive météorologique quelles sont les situations passées qui présentent les plus grandes similitudes. La connaissance des précipitations observées lors de ces situations analogues permet de fournir une estimation probabiliste des pluies attendues pour la situation cible. La caractérisation des situations cibles peut se faire soit via des données directement mesurées ou analysées (nous sommes alors dans un cadre de simulation), soit via les prévisions d'un modèle NWP (nous sommes alors dans un cadre de prévision).

La principale motivation de ce travail a été la récente apparition de longues archives de réanalyses météorologiques. Ces nouvelles archives garantissent d'une part une certaine homogénéité dans l'assimilation et l'analyse des données mesurées, et permettent d'autre part d'avoir accès à l'ensemble des variables atmosphériques. En utilisant un sous-ensemble des réanalyses NCEP/NCAR, nous avons donc recherché quelle était la manière la plus informative, vis à vis des précipitations, de définir une analogie météorologique.

Avant de nous atteler à l'essai de différents algorithmes d'analogie, nous avons dû définir les critères selon lesquels nous allons décider qu'un algorithme est plus performant qu'un autre. Cette étape préliminaire fut l'occasion d'explicitier les bases sur lesquelles repose le critère choisi, le *CRPS*, et d'en exprimer une décomposition en terme de *Finesse* et de *Justesse*.

Ensuite, la recherche des algorithmes les plus performants a été menée dans un contexte de simulation, les situations cibles étant caractérisées par les réanalyses NCEP/NCAR. Cette étape d'optimisation a ainsi montré dans un premier temps que :

- L'analogie à privilégier est celle portant sur la circulation synoptique, via la comparaison de champs de hauteurs géopotentielles par l'intermédiaire d'un critère de forme, et non pas d'un critère de distance moyenne. Il n'est pas nécessaire de considérer un nombre important de champs, deux apparaissant comme suffisants à conditions qu'ils soient considérés à l'intérieur de la période de cumul des précipitations.

- Le domaine géographique sur lequel l'analogie doit être considéré est d'extension limitée, de l'ordre de $20 \times 10^\circ$. Il est intéressant de déplacer cette fenêtre si l'on considère des groupements pluviométriques trop éloignés les uns des autres (≈ 300 km), mais nous avons montré qu'elle n'avait pas à être redéfinie en fonction de chaque situation cible, en considérant par exemple une classification en type de temps.

Cette analogie sur la circulation synoptique, que nous appelons analogie de premier ordre, affiche cependant un plafonnement dans ses possibilités, quel que soit le degré de raffinement appliqué à sa définition. L'information contenue dans la circulation synoptique explique une grande partie des précipitations (score moyen de performance de l'ordre de 38% par rapport à la climatologie), mais est insuffisante pour tout expliquer.

Pour aller plus loin, nous avons introduit une analogie supplémentaire : l'analogie de second ordre. Parmi les situations retenues sur la base de l'analogie de premier ordre, on effectue une seconde sélection, fondée sur la comparaison, via un critère de distance moyenne, de variables d'humidité (combinaison de l'eau précipitable et de l'humidité relative). La fenêtre d'analogie de cette seconde sélection est de taille bien plus réduite que celle utilisée lors de la première sélection : l'analogie de second ordre est une analogie locale.

L'utilisation d'algorithme d'analogie à deux sélections permet ainsi d'obtenir des performances moyennes de l'ordre de 43% par rapport à la climatologie. Or, il est apparu que plus on remonte le temps dans les réanalyses, moins les données d'humidités sont fiables. Ces performances devraient donc encore augmenter au fur et à mesure de la constitution d'archives fiables.

Une fois que les algorithmes d'analogie ont été définis, une étape de validation a permis de montrer la robustesse des résultats obtenus. Elle fut aussi l'occasion d'examiner de plus près, en prenant le point de vue d'un utilisateur de prévisions, quelles étaient les performances de l'adaptation par analogie en prévision de précipitation.

En examinant par ailleurs la sensibilité de nos algorithmes à l'archive météorologique employée, nous avons montré le rôle crucial que celle-ci jouait dans l'importance des performances atteintes. Il est ainsi primordial de disposer d'une archive homogène dans le temps et d'utiliser, en prévision, une caractérisation des situations cibles reposant sur des modèles similaires à ceux utilisés pour construire l'archive.

Enfin, grâce à la mise en place d'une chaîne opérationnelle de prévision au LTHE, nous avons été en mesure d'appliquer ces algorithmes d'analogie dans un contexte de prévision opérationnelle, basée sur les sorties du modèle GFS du NCEP. Cette application, qui inclut une étude de cas sur les précipitations extrêmes qui ont eut lieu les 8 et 9 septembre 2002 dans la région des Cévennes, a permis de constater que :

- les algorithmes d'analogie à deux sélections permettent effectivement d'améliorer les performances des algorithmes à une seule sélection, même dans un contexte opérationnel,
- mais seulement sur les premières échéances (J, J+1 et J+2), et de façon moindre que ce que laissait espérer les résultats obtenus en prévision parfaite.

Les incertitudes importantes qui entachent la prévision des variables d'humidité limitent donc les performances atteintes. Au delà de J+2, ces variables ne sont plus assez fiables pour permettre

d'améliorer une analogie de premier ordre. Les algorithmes à une et à deux sélections sont donc complémentaires l'un à l'autre :

- la première sélection est robuste et peut s'appliquer jusqu'à $J+5$ ou $J+6$, mais présente une finesse souvent insuffisante,
- tandis que la seconde sélection permet d'affiner notablement cette première sélection, mais se limite à J , $J+1$ et $J+2$ et risque de présenter une sensibilité accrue aux incertitudes de la modélisation numérique.

Ce travail a donc permis une amélioration significative des techniques de recherche d'analogie appliquées à la prévision des précipitations, mais nous permet surtout de mieux comprendre quelles sont les analogies les plus informatives, et pourquoi elles le sont. Nous pensons par ailleurs avoir fourni les éléments nécessaires pour que de potentiels utilisateurs puissent appréhender ce qu'une adaptation par analogie est susceptible de leur apporter comme information. On pourrait en outre envisager d'utiliser la base de données PLUVIO pour proposer un *pavage* complet de la France et intéresser ainsi le plus grand nombre d'utilisateurs.

Un autre aspect intéressant dans l'utilisation de l'adaptation par analogie est sa relative indépendance vis à vis de la paramétrisation physique des modèles numériques. Ces méthodes fournissent une référence qui permet de comparer différents modèles, aussi bien au niveau de la prévisions des précipitations que de celle de la circulation synoptique. L'adaptation, en générant les précipitations associées à une circulation donnée, joue ainsi un rôle d'amplificateur qui permet d'appréhender les effets des incertitudes qui entachent les prévisions synoptiques.

Ces techniques ne vont d'ailleurs sûrement pas en rester là. En effet, l'apparition de nouveaux produits météorologiques ouvre de motivantes perspectives :

- Ainsi, le tout récent achèvement des réanalyses ERA-40 du CEPMMT, disponibles à une résolution de $1 \times 1^\circ$, permet d'envisager une nette amélioration des algorithmes à deux sélections, ceux-ci étant particulièrement sensibles à la résolution des données d'humidité, comme nous l'avons vu au cours de ce travail. On attend par contre peu de profits notables sur l'analogie de premier ordre.
- En outre, l'utilisation des prévisions d'ensemble pourrait s'avérer très intéressante pour augmenter les échéances de prévisions. Au lieu de rechercher ainsi 50 ou 30 situations analogues à une seule prévision de la situation cible, on pourrait ainsi envisager de rechercher seulement 5 ou 10 analogues à chacune des traces de la prévision d'ensemble. De telles prévisions risquent éventuellement de présenter une dispersion trop importante pour être effectivement informatives, mais permettent d'éviter de produire des prévisions de précipitations qui se rapportent à des situations cibles qui n'existeront jamais réellement...

Si l'on se place du point de vue d'une utilisation hydrologique des prévisions par adaptation, il est fort dommage que le pas de temps des cumuls de précipitations prévus soit journalier. Cette limitation est principalement due à la disponibilité des données de pluies, mais au regard de la taille des bassins versants qui nous intéressent ici (500 à 2000 km²), de telles prévisions ne peuvent pas être intégrées telles quelles dans un modèle hydrologique. Elle présentent certes

un intérêt certain dans le cadre d'une mise en vigilance ou d'une mise en alerte, même si leur résolution temporelle n'est pas adaptée directement à la dynamique des bassins méditerranéens.

Deux pistes sont actuellement explorées au LTHE pour tenter de remédier à ce problème. La première consiste à essayer de fractionner les cumuls de pluie journaliers en deux cumuls de 12 heures, grâce à l'utilisation régionale de quelques postes témoins ayant fonctionnés à un pas de temps infra-journalier. L'autre piste consiste à créer des scénarii horaires de pluie, qui respectent d'une part la dynamique des précipitations propres à un lieu donné, et d'autre part la distribution des totaux journaliers prévus par analogie. L'application commune de ces deux solutions pourrait, du moins nous l'espérons, apporter une solution satisfaisante au problème de l'adaptation des prévisions de précipitation aux besoins des hydrologues.

Enfin, il serait faux de croire que ces techniques d'analogie ne s'appliquent qu'à des problèmes de prévisions. Comme nous l'avons déjà évoqué, elles peuvent servir à faire apparaître les mécanismes propres à certains types de situations, mais elles ont aussi leur place dans le cadre de la prévision climatique. Les modèles climatiques ayant une résolution assez grossière, afin d'être capable de simuler de longues périodes, les variables locales telles que les précipitations sont de faible qualité. L'application de l'analogie permet ainsi de produire des champs de précipitations réalistes à partir des prévisions des variables de grandes échelle, ainsi que l'a par exemple fait Martin (1995). La question qui se pose alors est de savoir quelle est la pertinence de l'utilisation d'une archive météorologique du climat actuel pour simuler un climat futur. La réponse actuelle serait qu'un changement du climat ne se traduirait pas par l'apparition de nouvelles situations, mais plutôt par un changement dans la fréquence d'occurrence de situations qui ont déjà été observées. Dans ce cas, l'application de l'analogie est légitime et constitue sûrement l'un des meilleurs exemples de la complémentarité qui existe entre la modélisation numérique et les techniques d'adaptation statistique.

Références bibliographiques

- Anderson J., H. Van den Dool, A. Barnston, W. Chen, W. Stren et J. Ploshay** (1999). Present-day capabilities of numerical and statistical models for atmospheric extratropical seasonal simulation and prediction. *Bulletin of the American Meteorological Society* **80**(7), pp. 1349–1361.
- Antolik M. S.** (2000). An overview of the National Weather Service’s centralized statistical quantitative precipitation forecasts. *Journal of Hydrology* **239**, pp. 306–337.
- Atger F.** (1999). The skill of ensemble prediction systems. *Monthly Weather Review* **127**, pp. 1941–1953.
- Atger F.** (2000). La prévision du temps à moyenne échéance en France. *La Météorologie* **8**(30), pp. 61–86.
- Bardossy A., L. Duckstein et I. Bogardi** (1995). Fuzzy rule-based classification of atmospheric circulations patterns. *International Journal of Climatology* **15**(10), pp. 1087–1097.
- Barnett T. P. et R. W. Preisendorfer** (1978). Multifield analog prediction of short-term fluctuations using a climate state vector. *Journal of the Atmospheric Sciences* **35**(10), pp. 1771–1787.
- Barnston A. G. et T. M. Smith** (1996). Specification and prediction of global surface temperature and precipitation from global SST using CCA. *Journal of Climate* **9**, pp. 2660–2697.
- Barry R. G. et A. H. Perry** (1973). *Synoptic climatology : methods and applications* (1ère édition ed.). Londres, R.U. : Methuen & Co Ltd, pp. 555.
- Bois Ph. et Ch. Obled** (1976). Prévision des avalanches par des méthodes statistiques : aspects méthodologiques et opérationnels. *La Houille Blanche* (6/7), pp. 509–531.
- Bois Ph. et Ch. Obled** (2001). *Introduction au traitement de données en hydrologie* (édition révisée ed.). Ecole Nationale Supérieure d’Hydraulique et de Mécanique de Grenoble : Institut National Polytechnique de Grenoble, pp. 230.
- Bois Ph., Ch. Obled et J. Thalamy** (1981). Etude des liaisons entre champs de pressions et températures sur l’Europe avec la durée d’insolation : application à la possibilité de prévision de l’insolation. In PIRDES COMES (Ed.), *C. R. du colloque Météorologie de l’Energie Solaire*, Toulouse, France, pp. 145–177.
- Bontron G., A. Djerboua et Ch. Obled** (2002). Sélection de situations météorologiques analogues : applications en prévision opérationnelle de précipitations et en évolution climatique. *La Houille Blanche* (8), pp. 46–51.

- Bontron G., G. Ménez, D. Duband et J.-N. Gautier** (1999). Application de la méthode du gradex à de grands bassins versants : cas de la Loire au bec d'allier (32 000 km²). *La Houille Blanche* (6), pp. 29–36.
- Bouttier F.** (1994). *Sur la prévision de la qualité des prévisions météorologiques*. Thèse de doctorat, Université Paul Sabatier.
- Brier G. W.** (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review* **78**(1), pp. 1–3.
- Bross I. D.** (1953). *Design for decision*, pp. 46–53. New York : The Macmillian Co.
- Brown T. A.** (1974). Admissible scoring systems for continuous distributions. Manuscript P-5235, The Rand Corporation, Santa Monica, CA.
- Buizza R. J., P. Tribbia, F. Molteni et T. N. Palmer** (1993). Computation of optimal unstable structures for a numerical weather prediction model. *Tellus* **45A**, pp. 388–407.
- Candille G.** (2003). *Validation des systèmes de prévisions météorologiques probabilistes*. Thèse de doctorat, Université Pierre et Marie Curie - Paris 6.
- Clement R.** (1976). Répartition géographique des précipitations d'hiver sur les Alpes françaises. Note interne 374, Etablissement d'études et de recherches météorologiques.
- Coiffier J.** (2000). Un demi-siècle de prévision numérique du temps. *La Météorologie* **8**(30), pp. 11–31.
- Delrieu G. et al.** (2003). The catastrophic flash-flood event of 8-9 September 2002 in the Gard region, France : a first case study for the Cévennes-Vivarais mediterranean hydro-meteorological observatory. *Journal of Hydrometeorology* **soumis**.
- Déqué M.** (2003). La prévision numérique à l'échelle saisonnière : que sait-on faire et que peut-on espérer ? *La Météorologie* **8**(41), pp. 20–29.
- Draper N. R. et H. Smith** (1981). *Applied regression analysis*. New York : Wiley, pp. 709.
- Duband D.** (1970). *Reconnaissance dynamique de la forme des situations météorologiques. Application à la prévision quantitative des précipitations*. Thèse de 3ème cycle, Faculté des sciences de Paris.
- Duband D.** (1971). Prévision quantitative des précipitations en automne. rapport interne, EDF-DTG.
- Duband D.** (1974). Reconnaissance dynamique de la forme des situations météorologiques. Application à la prévision quantitative des précipitations. In Société Hydrotechnique de France (Ed.), *XIIIèmes journées de l'Hydraulique*, Paris, France.
- Duband D.** (1980). Dynamic selection of analogue flow patterns to produce quantitative precipitation forecasts. In WMO (Ed.), *WMO Symposium on probabilistic and statistical method in weather forecasting*, Nice, France, 8-12 septembre.
- Duband D.** (1981). Prévision spatiale des hauteurs de précipitations journalières. *La Météorologie* (7/8), pp. 487–494.
- Ducrocq V., D. Ricard, J.-P. Lafore et F. Orain** (2002). Storm-scale numerical rainfall prediction for five precipitating events over France : on the importance of the initial humidity field. *Weather and Forecasting* **17**(6), pp. 1236–1256.

- Epstein E. S.** (1969). A scoring system for probability forecasts of ranked categories. *Journal of Applied Meteorology* **8**, pp. 985–987.
- Epstein E. S. et A. H. Murphy** (1965). A note on the attributes of probabilistic prediction and the probability score. *Journal of Applied Meteorology* **4**, pp. 297–299.
- Fehlmann R., C. Quadri et H. C. Davies** (2000). An alpine rainstorm : sensitivity to the mesoscale upper-level structure. *Weather and Forecasting* **15**, pp. 4–28.
- Foehn P., W. Good, Ph. Bois et Ch. Obled** (1977). Evaluation and comparison of statistical and conventional methods to forecast avalanche hazard. *Journal of Glaciology* **19**(11), pp. 375–387.
- Gibson J. K., P. Kalberg, S. Uppala, A. Noumura, A. Hernandez et E. Serrano** (1997). ERA description. ECMWF re-Analysis project report series 1, ECMWF, Reading, UK.
- Glahn H. R. et D. A. Lowry** (1972). The use of model output statistics (MOS) in objective weather forecasting. *Journal of Applied Meteorology* **11**, pp. 1203–1211.
- Golding B. W.** (2000). Quantitative precipitation forecasting in the UK. *Journal of Hydrology* **239**, pp. 286–305.
- Guilbaud S.** (1997). *Prévision quantitative des précipitations journalières par une méthode statistico-dynamique de recherche d'analogues : application à des bassins du pourtour méditerranéen*. Thèse de doctorat, Institut National Polytechnique de Grenoble.
- Guilbaud S. et Ch. Obled** (1998). L'approche par analogues en prévision météorologique. *La Météorologie* **8**(24), pp. 21–35.
- Hay L. E., G. J. McCabe, D. M. Wolock et M. A. Ayers** (1991). Simulation of precipitation by weather type analysis. *Water Resources Research* **27**(4), pp. 493–501.
- Hersbach H.** (2000). Decomposition of the continuous ranked probability score for ensemble prediction systems. *Weather and Forecasting* **15**, pp. 559–570.
- Huet P., X. Martin, J. L. Prime, P. Foin, C. Laurain et P. Cannard** (2003). Retour d'expérience des crues de septembre 2002 dans les départements du Gard, de l'Hérault, du Vaucluse, des Bouches du Rhône, de l'Ardèche et de la Drome. Rapport de l'inspection générale de l'environnement, Ministère de l'Ecologie et du Développement Durable, République Française.
- Juvanon du Vachat R.** (1994). Panorama de la prévision numérique à échelle fine ou à domaine limité. *La Météorologie* **8**(6), pp. 31–47.
- Kalnay E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, M. Gandin, M. Iredell, S. Saha, G. White, J. Woolen, Y. Zhu, M. Chelliah et W. Ebisuzaki** (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society* **77**(3), pp. 437–471.
- Karl T. R., M. E. Schlesinger et W. C. Wang** (1989). A method of relating general circulation model simulated climate to the observed local climate. Part I : central tendencies and dispersion. In A. M. Society (Ed.), *Sixth conference on applied climatology*, pp. 188–196.
- Katz R. W., A. H. Murphy et R. L. Winkler** (1982). Assessing the value of frost forecasts to orchardists : a dynamic decision-making approach. *Journal of Applied Meteorology* **21**, pp. 518–531.
- Krzysztofowicz R.** (1983). Why should a forecaster and a decision maker use bayes theorem. *Water Resources Research* **19**(2), pp. 327–336.

- Krzysztofowicz R.** (1992). Bayesian correlation score : a utilitarian measure of forecast skill. *Monthly Weather Review* **120**, pp. 208–219.
- Krzysztofowicz R.** (1998). Probabilistic hydro-meteorological forecasts : toward a new era in operational forecasting. *Bulletin of the American Meteorological Society* **79**(2), pp. 243–251.
- Lebel T., G. Bastin, Ch. Obled et D. Creutin** (1987). On the accuracy of areal rainfall estimation : a case study. *Water Resources Research* **23**(11), pp. 2123–2134.
- Livezey R. E. et A. G. Barnston** (1988). An operational multifield analog/antianalog prediction system for United States seasonal temperatures. Part I : System design and winter experiments. *Journal of Geophysical Research* **93**, pp. 10.953–10.974.
- Livezey R. E. et A. G. Barnston** (1989). An operational multifield analog/antianalog prediction system for United States seasonal temperatures. Part II : Spring, summer, fall and intermediate 3-month period experiments. *Journal of Climate* **2**, pp. 513–541.
- Lorenz E. N.** (1956). Empirical orthogonal functions and statistical weather prediction. Science Report No. 1, Contract AF 19(609)-1566, Dept. of Meteorology, MIT.
- Lorenz E. N.** (1963). Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences* **20**, pp. 130–141.
- Lorenz E. N.** (1969). Atmospheric predictability as revealed by naturally occurring analogues. *Journal of the Atmospheric Sciences* **26**, pp. 636–646.
- Lorenz E. N.** (1993). *The essence of chaos*. Seattle, USA : University of Washington Press, pp. 227.
- Martin E.** (1995). *Modélisation de la climatologie nivale des Alpes française, application des techniques de régionalisation à l'étude de l'impact d'un changement climatique sur l'enneigement*. Thèse de doctorat, Université Paul Sabatier.
- Massé P.** (1946). *Les réserves et la régulation de l'avenir dans la vie économique*. Paris, France : Hermann & Cie, pp. 138.
- Massacand A. C., H. Wernli et H. C. Davies** (1998). Heavy precipitation on the alpine southside : an upper-level precursor. *Geophysical Research Letter* **25**, pp. 1435–1438.
- Matheson J. E. et R. L. Winkler** (1976). Scoring rules for continuous probability distributions. *Management Science* **22**(10), pp. 1087–1096.
- Moch O.** (2002). «Donnez-moi un problème insoluble!». *Atmosphériques* (juillet), pp. 39.
- Murphy A. H.** (1971). A note on the ranked probability score. *Journal of Applied Meteorology* **10**, pp. 155–156.
- Murphy A. H.** (1973). A new vector partition of the probability score. *Journal of Applied Meteorology* **12**, pp. 595–600.
- Murphy A. H. et B. G. Brown** (1983). Forecast terminology : composition and interpretation of public weather forecasts. *Bulletin of the American Meteorological Society* **64**, pp. 13–22.
- Murphy A. H. et H. Daan** (1985). Forecast evaluation. In A. H. Murphy and R. W. Katz (Eds.), *Probability, statistics, and decision making in the atmospheric sciences*, pp. 379–437. Westview press.
- Murphy A. H. et E. S. Epstein** (1967). Verification of probabilistic predictions : a brief review. *Journal of Applied Meteorology* **6**, pp. 748–755.

- Murray F. W.** (1967). On the computation of saturation vapor pressure. *Journal of Applied Meteorology* **6**, pp. 203–204.
- Obled Ch. et Ph. Bois** (1982). Techniques d'analyse discriminante et de sélection d'analogues pour la prévision de l'insolation à courte échéance. In PIRDES COMES (Ed.), *C. R. du colloque Météorologie de l'Energie Solaire*, Toulouse, France, pp. 154–175.
- Obled Ch., G. Bontron et R. Garçon** (2003). Quantitative precipitation forecasts : a statistical adaptation of model outputs through an analogues sorting approach. *Atmospheric Research* **63**, pp. 303–324.
- Obled Ch. et W. Good** (1980). Recent developments in Avalanche Forecasting by statistical techniques : a methodological review and some applications to the parsenn area (Davos-Switzerland). *Journal of Glaciology* **25**(92), pp. 315–346.
- Orlanski I.** (1975). A rational subdivision of scales for atmospheric processes. *Bulletin of the American Meteorological Society* **56**, pp. 527–530.
- Pailleux J., J.-F. Geleyn et E. Legrand** (2000). La prévision numérique du temps avec les modèles Arpège et Aladin. *La Météorologie* **8**(30), pp. 32–60.
- Pham D. T.** (2001). Notion de probabilité pour l'ICA. In Ch. Jutten and A. Guérin-Dugué (Eds.), *Actes de l'école de printemps : de la séparation des sources à l'analyse en composantes indépendantes : méthodes, algorithmes et applications*, Villard-de-Lans, France, 2 au 4 mai 2001, pp. 1–42.
- Poccard I.** (2000). *Etude diagnostique de nouvelles données climatiques : les réanalyses : exemple d'application aux précipitations en Afrique Tropicale*. Thèse de doctorat, Université de Bourgogne.
- Potts J. M., C. K. Folland, I. T. Jolliffe et D. Sexton** (1996). Revised LEPS scores for assessing climate model simulations and long-range forecasts. *Journal of Climate* **9**, pp. 34–53.
- Rew R. K. et G. P. Davis** (1997). Unidata's netCDF interface for data access : status and plans. In American Meteorological Society (Ed.), *Thirteenth International conference on interactive information and processing systems for meteorology, oceanography and hydrology*, Anaheim, CA, USA, février, pp. 154–175.
- Roebber P. J. et G. W. Reuter** (2002). The sensitivity of precipitation to circulation details. Part II : mesoscale modeling. *Monthly Weather Review* **130**, pp. 3–23.
- Rousseau D.** (1980). A new skill score for the evaluation of yes/no forecasts. In WMO (Ed.), *WMO Symposium on probabilistic and statistical method in weather forecasting*, Nice, France, 8-12 septembre.
- Ruostenoja K.** (1988). Factors affecting the occurrence and lifetime of 500 mb height analogues : a case study based on large amount of data. *Monthly Weather Review* **116**, pp. 368–376.
- Sanders F.** (1958). The evaluation of subjective probability forecasts. Scientific report 5, Cambridge, Massachusetts Institute of Technology, Departement of Meteorology, Contract AF 19(604)-1305.
- Sanders F.** (1963). On subjective probability forecasting. *Journal of Applied Meteorology* **2**(2), pp. 191–201.

- Sharma A.** (2000). Seasonal to interannual rainfall probabilistic forecasts for improved water supply management : Part 3 - A nonparametric probabilistic forecast model. *Journal of Hydrology* **239**, pp. 249–258.
- Stanski H. R., L. J. Wilson et W. R. Burrows** (1989). Survey of common verification methods. Meteorology research report 89-5, Service de l'environnement atmosphérique, Downsview, Canada.
- Stehlik J. et A. Bardossy** (2002). Multivariate stochastic downscaling model for generating daily precipitation series based on atmospheric circulation. *Journal of Hydrology* **256**, pp. 120–141.
- Stuart A.** (1982). On the economic value of probability of precipitation forecasts in Canada. *Journal of Applied Meteorology* **21**, pp. 495–498.
- Talagrand O., R. Vautard et B. Strauss** (1997). Evaluation of probabilistic prediction systems. In *C. R. Atelier Predictability*, Reading, GB, Octobre, 1997, 1-25. Centre Européen pour les Prévisions à Moyen Terme.
- Teweless J. et H. Wobus** (1954). Verification of prognosis charts. *Bulletin of the American Meteorological Society* **35**(10), pp. 455–463.
- Thillet J.-J.** (1997). *La météo de montagne. Collection : Les guides du Club Alpin Français*. Paris : Seuil, pp. 189.
- Toth Z. et E. Kalnay** (1993). Ensemble forecasting at NMC : the generation of perturbations. *Bulletin of the American Meteorological Society* **74**, pp. 2317–2230.
- Triplet J. P. et G. Roche** (1986). *Météorologie générale* (3ème ed.). Météo-France, pp. 317.
- Uppala S., J. K. Gibson, M. Fiorino, A. Hernandez, P. Kalberg, X. Li, K. Onogi et S. Saarinen** (1999). ECMWF second generation reanalysis : ERA 40. In S. Organisation Météorologique Mondiale, Genève (Ed.), *2nd International conference of Reanalyses*, Reading, UK, 23-27 août, pp. 154–175.
- Van Den Dool H. M.** (1989). A new look at weather forecasting through analogues. *Monthly Weather Review* **17**, pp. 2230–2247.
- Van Den Dool H. M.** (1994). Searching for analogues, how long must we wait ? *Tellus* **46A**, pp. 314–324.
- Ward M. N. et C. K. Folland** (1991). Prediction of seasonal rainfall in the north nordeste of Brazil using eigenvectors of sea-surface temperature. *International Journal of Climatology* **11**, pp. 711–743.
- Wilks D. S.** (1995). *Statistical methods in the atmospheric sciences*. San Diego, CA, USA : Academic Press, pp. 467.
- Wilson L. J., B. Strauss et A. Lanzinger** (1996). On the verification of weather element forecasts from an ensemble prediction system. In *Preprints, 15th Conference on Weather Analysis and Forecasting*, Norfolk, VA, pp. J83–J86. American Meteorological Society, Boston, MA.
- Woodcock F.** (1980). On the use of analogues to improve regression forecasts. *Monthly Weather Review* **108**, pp. 292–297.
- Zeng X., R. A. Pielke et R. Eykholt** (1993). Chaos theory and its application to the atmosphere. *Bulletin of the American Meteorological Society* **74**, pp. 631–644.
- Zevin S. F.** (1994). Steps toward an integrated approach to hydrometeorological forecasting services. *Bulletin of the American Meteorological Society* **75**(7), pp. 1267–1276.

- Zhu Y., Z. Yyengar, S. Toth, M. Tracton et T. Marchok** (1996). Objective evaluation of the NCEP global ensemble forecasting system. In A. M. Society (Ed.), *15th Conf. on Weather Analysis and Forecasting*, Norfolk, VA, pp. J79–J82.

Annexes

L'interface de visualisation des prévisions : RainFAST

A.1 Introduction

Un prototype opérationnel d'adaptation du modèle météorologique GFS (cf. chapitre 9), par recherche de situations analogues, a été mis en place au LTHE. Il fournit tous les jours, de façon routinière, des prévisions de précipitations journalières pour les 7 jours à venir, et ceci sur l'ensemble des groupements pluviométriques décrit dans ce mémoire. Les prévisions sont ensuite archivées sous la formes de fichiers ASCII indépendants les uns des autres.

Afin de permettre une visualisation sous forme graphique de ces prévisions, nous avons développé l'interface RainFAST (Rainfall Forecasts by Analog Sorting Techniques), laquelle permet :

- d'une part de consulter les prévisions de précipitations émises sur chacun des groupements par les algorithmes d'adaptation,
- et d'autre part d'avoir accès aux cartes météorologiques de la situation candidate (données fournies par le modèle GFS) et des situations reconnues comme analogues (réanalyses NCEP/NCAR).

RainFAST fonctionne sous IDL. Les fichiers sources sont de simples fichiers ASCII qui peuvent évidemment être édités et modifiés. Il est cependant conseillé, en cas de modifications, de faire celles-ci sur des copies des fichiers originaux, afin de laisser ces derniers inchangés.

⚠ | *Les différentes illustrations présentées ici sont issues de captures d'écran réalisées sous Linux. Leur aspect peut donc légèrement différer de ce qui apparaît sur un ordinateur fonctionnant sous un autre système d'exploitation.*

A.2 Mise en route

L'interface de visualisation se lance par l'intermédiaire d'IDL de la manière suivante :

1. se placer dans le répertoire contenant le fichier `RainFAST.pro` ;
2. lancer IDL ;
3. dans la ligne de commande IDL, taper `.r RainFAST` afin de compiler le programme ;
4. dans la ligne de commande IDL, taper `RainFAST` afin de lancer l'exécution du programme.

Une boîte de dialogue (fig. A.1) s'affiche alors. Elle permet de choisir la date d'émission de la prévision que l'on veut visualiser (il est ainsi possible de visualiser d'anciennes prévisions archivées). Ce choix se fait dans la fenêtre de saisie située sous le texte **Date de la prévision :**, sous la forme suivante : *jj/mm/aaaa* où *jj* correspond au jour, *mm* au mois et *aaaa* à l'année de la date désirée. Ainsi, pour obtenir la prévision émise le 3 octobre 2003, il faut entrer 03/10/2003, comme présenté sur la figure A.1.



FIG. A.1: Boîte de dialogue d'ouverture de RainFAST

⚠ Pour que la date désirée soit bien prise en compte, il ne faut pas oublier de taper sur la touche *ENTREE* à la fin de la saisie.

Si une prévision a effectivement été archivée pour la date désirée, les boutons **Cartes de Prévisions** et **Champs Synoptiques** deviennent actifs. Ils permettent d'accéder respectivement à :

- l'affichage des prévisions de précipitations (**Cartes de Prévisions**),
- et à celui des données météorologiques de la situation cible et des situations reconnues comme analogues (**Champs Synoptiques**).

De même, les deux boutons disposés sous la fenêtre de saisie deviennent eux aussi actifs :

- : recule la date affichée d'une journée,
- : avance la date affichée d'une journée.

A.3 Affichage des prévisions de précipitations

La sélection du bouton **Cartes de Prévisions** depuis la boîte de dialogue d'ouverture de RainFAST permet d'entrer dans le mode de visualisation des prévisions de précipitations. L'écran de travail est alors tel que présenté par la figure A.2.

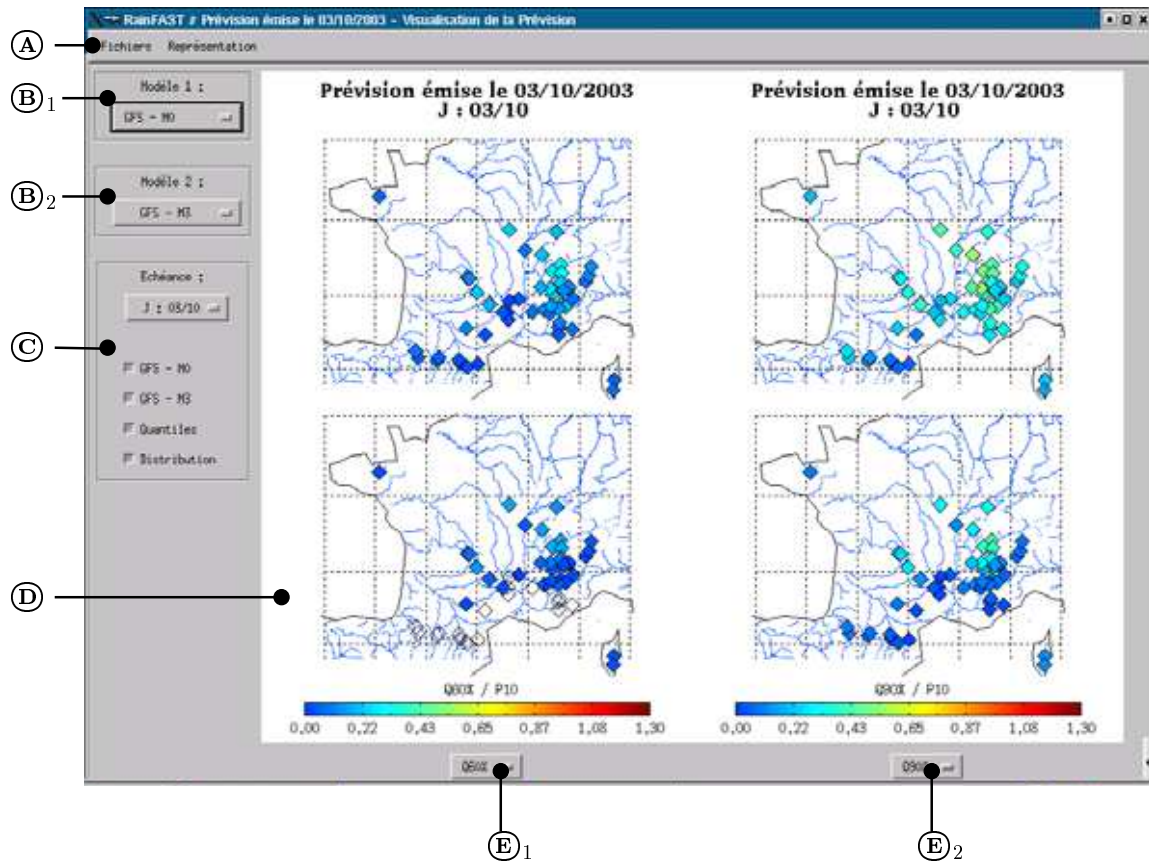


FIG. A.2: Ecran de travail en mode **Cartes de Prévisions**.

A.3.1 Les cartes de prévisions

La fenêtre graphique ① se décompose en 4 parties. Chacune d'elles cartographie les valeurs prises par l'un des quantiles de la prévision sur l'ensemble des groupements cibles, selon (i) un algorithme d'adaptation et (ii) une échéance de prévision.

Choix des algorithmes d'adaptation

RainFAST permet de comparer aisément les résultats de deux algorithmes d'adaptation différents, en affichant simultanément leurs résultats. Le choix de ces algorithmes se fait par l'intermédiaire des menus déroulant ②₁ et ②₂ :

- ②₁ concerne les deux cartes supérieures,
- ②₂ concerne les deux cartes inférieures.

Ces menus proposent actuellement les choix suivants (l'introduction ultérieure de nouveaux algorithmes pourrait amener à accroître cette liste) :

- GFS - M0 : adaptation du modèle GFS selon l'algorithme M0 (cf. chapitre 7),
- GFS - 2 champs : adaptation du modèle GFS selon l'algorithme M1, mais en utilisant une fenêtre d'analogie commune optimisée et non pas une fenêtre pour chaque regroupement (cf. chapitre 5),
- GFS - M3 : adaptation du modèle GFS selon l'algorithme M3,
- GFS - M1 : adaptation du modèle GFS selon l'algorithme M1.

A chaque action sur les menus \textcircled{B}_1 et \textcircled{B}_2 , la fenêtre graphique \textcircled{D} est mise à jour.

Choix de l'échéance de prévision

Le choix de l'échéance de prévision représentée s'effectue via le menu déroulant situé dans le cadre \textcircled{C} (figure A.3 : exemple d'une prévision émise le 03/10/2003). Ce choix affecte simultanément les quatre cartes affichées dans la fenêtre graphique \textcircled{D} .

Nous rappelons que l'échéance J correspond au cumul de précipitations attendu entre 6 hTU le jour de l'émission de la prévision et 6 hTU le lendemain. L'échéance J+1 correspond au cumul attendu entre 6 hTU le lendemain du jour de l'émission de la prévision et 6 hTU le jour suivant, et ainsi de suite ...

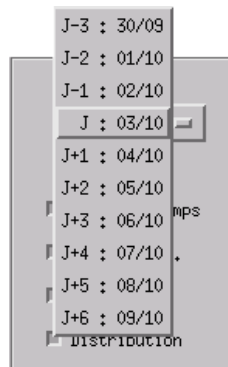


FIG. A.3: Menu déroulant du cadre \textcircled{C} : sélection de l'échéance de prévision (exemple dans le cas d'une prévision émise un 3 octobre).

L'échéance notée J-1 correspond à la prévision émise la veille du jour J pour le cumul de précipitation allant de 6hTU le jour J-1 à 6hTU le jour J. Ainsi, dans le cas de la visualisation d'une prévision émise un 3 octobre, cette prévision à J-1 correspond à la prévision pour le 2 octobre émise le 2 octobre à 0 hTU. De même, J-2 correspond à l'avant-veille du jour J et J-3 au jour encore précédent. Ces échéances ne sont disponibles que si les prévisions correspondantes ont été préalablement archivées.

Choix du quantile de prévision

Les menus déroulants \textcircled{E}_1 et \textcircled{E}_2 permettent de choisir les quantiles de prévision que l'on souhaite visualiser. \textcircled{E}_1 (respectivement \textcircled{E}_2) concerne les deux cartes situées dans la partie gauche (resp. la partie droite) de la fenêtre graphique \textcircled{D} . Les choix disponibles sont :

- Q20%, Q60% et Q90% qui correspondent respectivement au quantile au non-dépassement 20, 60 et 90 % de la prévision de précipitation,
- **max** qui correspond à la plus forte valeur de précipitation associée aux situations reconnues comme analogues.

⚠ | Rappelons que les valeurs de précipitations prévues sont des lames d'eau moyennes sur un ensemble de postes pluviométriques, cumulées sur 24 heures.

Les valeurs prévues peuvent être exprimées soit en fraction de la pluie décennale de chaque groupement (mode par défaut, tel que présenté dans la figure A.2) ou directement en hauteur de pluie mesurée. Le passage d'un mode de représentation à l'autre se fait via le menu **Représentation** de la barre de menu ④.

A.3.2 Détails par groupement

Si l'on déplace la souris au-dessus de l'un des groupements dans la fenêtre graphique ⑩, le nom de ce groupement apparaît dans le cadre ③, en dessous du menu déroulant permettant de sélectionner l'échéance de prévision affichée (figure A.4, exemple en cas de sélection du groupement **Arly**). En cliquant sur le groupement désiré, on fait apparaître une nouvelle fenêtre qui permet d'afficher le détail de la prévision pour ce groupement. En fonction des boutons activés dans le cadre ③, différentes informations sont disponibles.

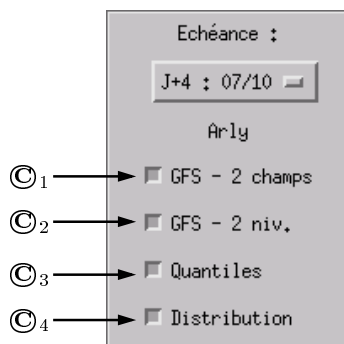


FIG. A.4: Cadre ③ lorsque la souris est positionnée au-dessus du groupement **Arly**.

La figure A.5 présente l'exemple d'une fenêtre de détail de prévision lorsque les boutons ③₂, ③₃ et ③₄ sont activés.

Les boutons ③₁ et ③₂, dont le libellé correspond aux noms des algorithmes sélectionnés via les menus déroulants ②₁ et ②₂, permettent de désigner quelles sont les prévisions que l'on souhaite détailler. Au moins l'un de ces boutons doit être activé ; si les deux sont activés simultanément, les détails des deux types de prévisions apparaissent l'un au dessus de l'autre.

Les boutons ③₃ et ③₄ permettent de désigner quel type de détail est désiré (de même que précédemment, au moins l'un de ces bouton doit être sélectionné) :

- ③₃, dénommé **Quantiles**, fournit l'évolution des quantiles prévus Q20%, Q60% et Q90% sur le groupement considéré, pour les 7 jours à venir, de J à J+6 (panneau de gauche de la

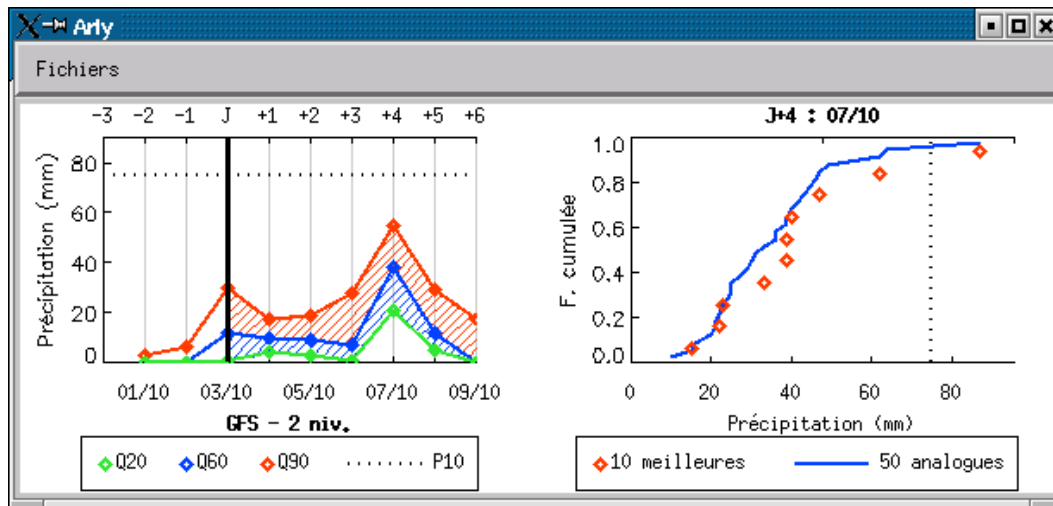


FIG. A.5: Exemple de fenêtre donnant le détail de la prévision pour le groupement Arly.

figure A.5). Si les prévisions sont disponibles, les valeurs prévues à échéance de J lors des trois jours précédents sont aussi rappelées (échéance J-3, J-2 et J-1).

- ©₄, dénommé **Distribution**, fournit la distribution des pluies prévues sur le groupement considéré, et ceci pour l'échéance préalablement désignée via le menu déroulant du cadre © (panneau de droite de la figure A.5) :
 - > la courbe pleine correspond à la distribution donnée par l'ensemble des situations analogues (dont le nombre varie en fonction de l'algorithme considéré),
 - > les losanges correspondent à la distribution des 10 meilleures analogues seulement.
 L'écart entre ces deux distributions fournit des indications sur la rareté de la situation météorologique.

Dans un cas comme dans l'autre, la courbe pointillée indique la valeur de la pluie décennale sur le groupement.

A.4 Visualisation des données météorologiques

Dans le mode de visualisation des données météorologiques, l'écran de travail est tel que présenté par la figure A.6. Ce mode permet de comparer les situations météorologiques reconnues comme analogues par les algorithmes d'adaptation (données issues des réanalyses NCEP/NCAR), soit entre elles, soit avec la situation météorologique cible, telle que prévue par le modèle GFS. Il est par ailleurs possible de comparer simultanément deux champs météorologiques différents.

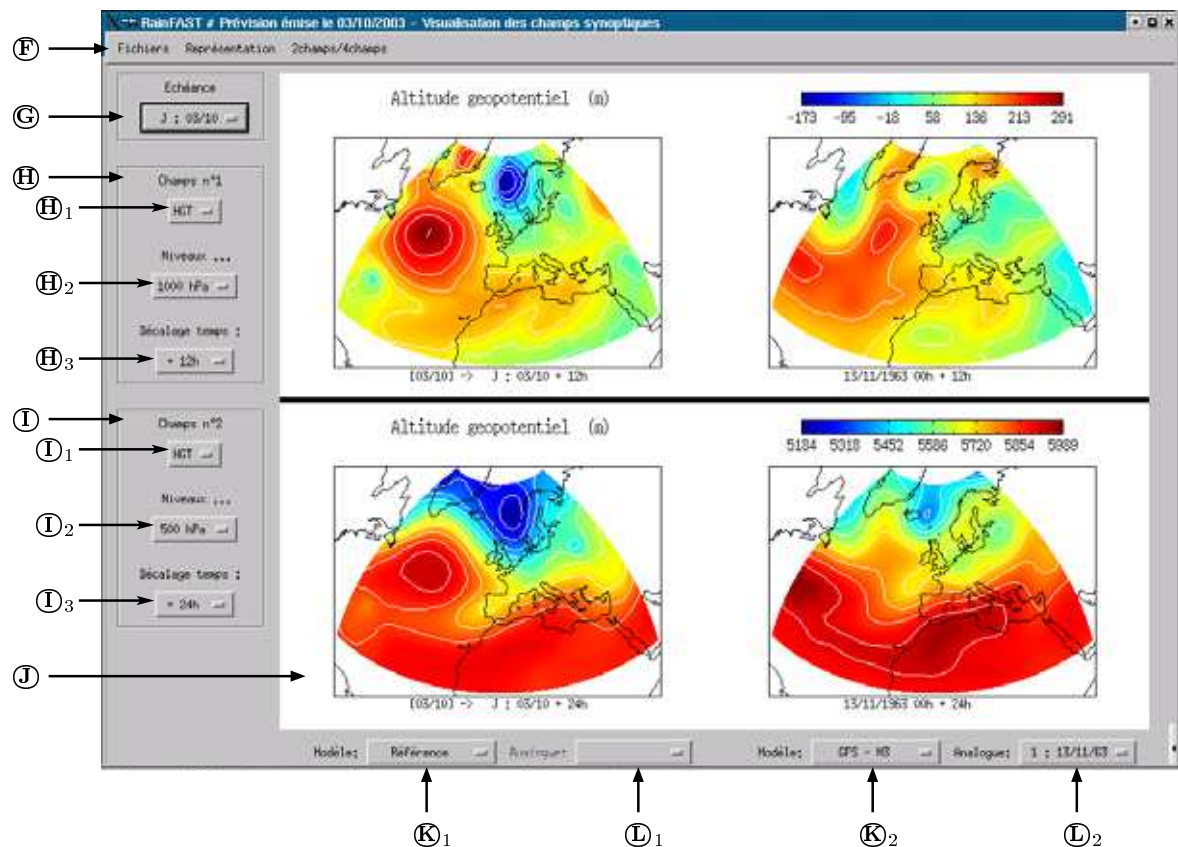


FIG. A.6: Ecran de travail en mode Champs synoptiques.

A.4.1 Sélection des données à comparer

Choix des situations

Par défaut, RainFAST présente d'une part la situation cible (partie gauche de la fenêtre graphique **J**), et d'autre part la meilleur analogue au sens de l'algorithme GFS - M3 pour cette situation cible (partie droite).

Les menus déroulant **K**₁ et **K**₂ permettent de sélectionner les algorithmes dont sont issues les situations présentées respectivement dans la partie gauche et la partie droite de la fenêtre graphique. Les options offertes sont :

- **Référence** : la situation affichée est la situation cible, telle que prévue par le modèle GFS. Le menu déroulant **L** associé est alors rendu inactif car nous ne disposons que d'une seule prévision de la situation cible.
- **GFS - M0, GFS - 2 champs, GFS - M3 et GFS - M1** : permet d'afficher les situations reconstruites comme analogues par l'un de ces algorithmes. Les données météorologiques sont alors issues des réanalyses NCEP/NCAR. Le choix de la situation analogue se fait via le menu déroulant **L** associé, qui liste les 20 meilleures analogues au sens de l'algorithme considéré.

Choix des champs météorologiques

Par défaut, les champs météorologiques représentés sont ceux utilisés par les algorithmes M1 et M2, à savoir les géopotentiels 1000 hPa +12h (panneau supérieur) et 500 hPa +24h (panneau inférieur).

Les cadres **Ⓜ** et **Ⓜ** permettent de modifier ces choix : **Ⓜ** (resp. **Ⓜ**) agit sur le panneau supérieur (resp. le panneau inférieur) de la fenêtre graphique **Ⓜ**. Nous n'illustrons ici que la manipulation du cadre **Ⓜ**, le cadre **Ⓜ** fonctionnant de manière parfaitement similaire.

Le choix du champs météorologique a lieu en 3 temps :

1. Avec le bouton **Ⓜ**₁, sélectionner la variable météorologique désirée (fig. A.7 a). Celles-ci sont désignées par leurs codes, suivant le tableau A.1.
2. Le bouton **Ⓜ**₂ se met alors automatiquement à jour pour ne proposer que les niveaux disponibles pour la variable sélectionnée. Choisir alors avec ce bouton le niveau désiré (fig. A.7 b).
3. Enfin, le bouton **Ⓜ**₃ permet de définir l'heure de mesure du champ choisi (+00h correspond à 0 hTU, soit 6 heures avant le début du cumul des précipitations).

La représentation graphique de la situation météorologique se met automatiquement à jour à chaque intervention sur l'un de ces boutons.

Code	Variable	Type	Niveaux (hPa)						heure (hTU)	
HGT	Altitude du géopotentiel	A	1000	850	700	500	300	200	00	12
TMP	Température	A		850		500		200	00	12
UGR	Vitesse du vent zonal	A		850		500		200	00	12
VGR	Vitesse du vent méridien	A		850		500		200	00	12
RHU	Humidité relative	B		850		500			00	12
VVE	ω (<i>i.e.</i> vitesse verticale en coordonnées « pression »)	B				500			00	12
PWA	Eau précipitable	B	cumul sur la colonne d'air						00	12

TAB. A.1: Variables météorologiques et niveaux géopotentiels disponibles.

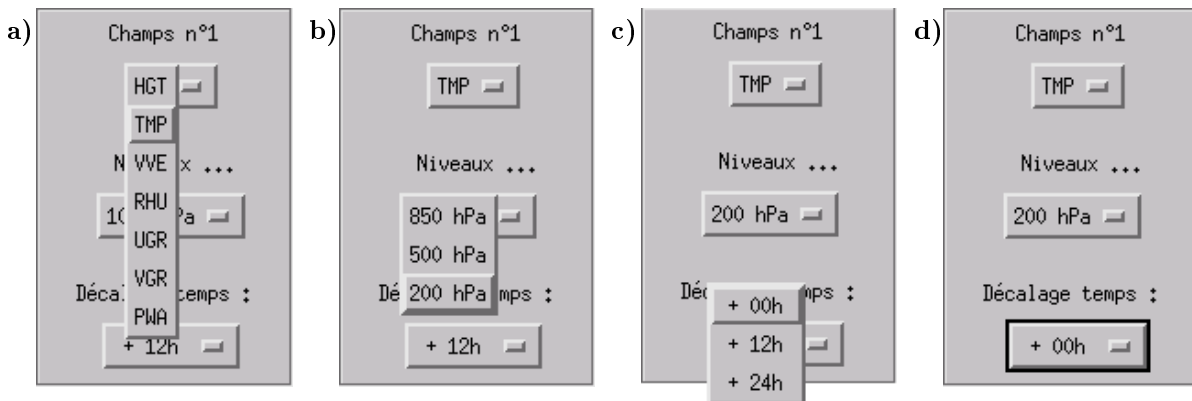


FIG. A.7: Sélection d'un champ météorologique via le cadre **Ⓜ**. **a)** sélection de la variable, **b)** sélection du niveau, **c)** sélection du décalage de l'heure de mesure, **d)** cadre **Ⓜ** après sélection.

Choix de l'échéance de prévision

Le menu déroulant ③ permet de sélectionner l'échéance de prévision, de manière analogue au menu du cadre ② présent dans le mode de visualisation des prévisions de précipitation. Son action agit sur les 4 cartes à la fois et met à jour les situations disponibles via les menus déroulants ①₁ et ①₂.

A.4.2 Options de représentation

Le menu déroulant **Représentation** de la barre de menu ⑤ propose plusieurs options permettant de modifier la représentation graphique des champs météorologiques (fig. A.8).



FIG. A.8: **a)** affichage du menu **Représentation**; **b)** affichage du sous-menu **Projection**.

Les 5 premières options permettent d'activer/désactiver certaines propriétés de l'affichage. Lorsqu'une option est précédée par (+) (respectivement par ()), celle-ci est activée (resp. désactivée). Ainsi, dans l'exemple de la figure A.8, seules les options **Interpolation** et **Surfaces** sont activées. Pour changer l'état d'activation d'une option, il suffit de cliquer dessus dans le menu **Représentation**. Ces options sont :

- **Interpolation** : Active l'interpolation bilinéaire des données afin de présenter une carte lissée. Si cette option est désactivée, chaque maille ($2.5^\circ \times 2.5^\circ$) est représentée par un carré de couleur unie. Cette option n'a d'effets que si l'option **Surfaces** est activée.
- **Grilles** : Permet d'afficher méridiens et parallèles sur la carte (espacement de 5°).
- **Surfaces** : Si cette option est activée, les champs sont représentés par une image (aplats de couleurs), sur laquelle sont sur-imposées des courbes de niveaux. Si elle est désactivée, le champ est seulement représenté par ses courbes de niveaux.
- **Domaine** : Permet d'afficher les limites du domaine d'analogie correspondant à l'algorithme sélectionné.
- **Zoom** : Permet de réduire le domaine géographique représenté.

L'option suivante, **Projection**, ouvre un sous-menu dans lequel plusieurs choix sont possibles (ces options ont vocation à être complétées en fonction des besoins) :

- **Orthographique** : Adopte une projection orthographique centrée sur 45° N.
- **Lon/Lat** : N'effectue pas de projection géographique. Les cartes sont tracées directement en fonction des longitudes et latitudes. Dans ce mode de représentation, les limites des surfaces continentales ne sont pas tracées.

A.5 Changement de date et de mode de visualisation

Changement de date d'émission de la prévision

Il est toujours possible de visualiser des prévisions émises pour une autre date, à condition que celles-ci aient été archivées.

Pour ce faire, dans le menu **Fichiers** de la barre de menu **(A)** ou **(F)**, selon le mode de visualisation courant, il faut sélectionner l'entrée **Changement date**. L'écran de travail disparaît alors et l'on retrouve la boîte de dialogue d'ouverture de RainFAST, dans laquelle on peut choisir une date d'émission différente (cf. A.2).

Changement de mode de visualisation

Pour passer d'un mode de visualisation à l'autre, RainFAST propose deux possibilités :

- Il est possible de retourner à la boîte de dialogue d'ouverture de RainFAST (par l'intermédiaire de l'entrée **Changement date** du menu **Fichiers** de la barre de menu **(A)** ou **(F)**). Dans cette boîte de dialogue d'ouverture, on clique alors sur le bouton **Champs Synoptiques** ou **Cartes de Prévisions** en fonction du mode de visualisation désiré. En procédant ainsi, il est possible de changer la date d'émission de la prévision.
- On peut aussi sélectionner directement l'entrée **Champs synoptiques** (resp. **Cartes**) du menu **Fichiers** de la barre de menu **(A)** (resp. **(F)**). L'écran de travail passe alors directement dans le mode de visualisation des champs météorologiques (resp. des prévisions de précipitations), mais la date d'émission de la prévision reste la même.

A.6 Exporter et imprimer les graphiques

A tout moment, il est possible d'exporter le contenu d'une fenêtre graphique (visualisation des prévision, des données météorologiques ou du détail de la prévision sur un bassin) dans un fichier au format Encapsulated PostScript :

1. dans le menu **Fichiers** de la barre de menu située en haut de la fenêtre dont on veut sauvegarder le contenu, choisir l'entrée **Exporter -> PS** ;
2. dans la boîte de dialogue qui s'affiche alors, choisir le chemin et le nom du fichier résultat ;
3. le fichier est créé !

Les fichiers au format Encapsulated PostScript (extension **.eps**) peuvent être facilement visualisés, imprimés et exportés vers d'autres formats via des logiciels à licence libre tels que *GSview* ou *Adobe Acrobat Reader*. C'est par ce biais qu'il est possible d'imprimer le contenu des différentes fenêtres graphiques.

THESE DE DOCTORAT DE L'INPG

Titre de l'ouvrage :

PREVISION QUANTITATIVE DES PRECIPITATIONS : ADAPTATION PROBABILISTE PAR RECHERCHE D'ANALOGUES.

**Utilisation des Réanalyses NCEP/NCAR
et application aux précipitations du quart Sud-Est
de la France**

Nom de l'auteur :

Guillaume BONTRON

Etablissement :

**Laboratoire d'étude
des Transferts en Hydrologie et Environnement
(LTHE, UMR 5564, CNRS-INPG-IRD-UJF)**

Résumé :

Ce mémoire étudie l'application du principe d'*analogie* à la prévision et à la simulation des précipitations. A partir de la caractérisation d'une situation météorologique cible, cette approche consiste à rechercher, dans une archive météorologique, les situations passées similaires. Les précipitations observées lors de ces situations analogues permettent ensuite de donner une estimation probabiliste des précipitations attendues pour la situation cible.

Après avoir défini les critères et les scores qui nous permettent de comparer deux prévisions probabilistes, nous avons cherché à identifier les algorithmes d'analogie les plus informatifs vis à vis des précipitations. Pour ce faire, nous nous sommes appuyé sur les réanalyses NCEP/NCAR. La richesse de cette archive nous a permis de définir des algorithmes d'analogie qui intègrent à la fois (i) l'information à grande échelle contenue dans la circulation synoptique, et (ii) l'information à échelle plus locale contenue dans les variables d'humidité.

Les différents paramètres de la méthode (choix des variables, domaine à prendre en compte, ...) ont été optimisés lors d'une phase de calibration. La validation des algorithmes obtenus a ensuite permis d'apprécier les performances atteintes en adoptant le point de vue d'un utilisateur, ainsi que d'étudier la sensibilité de ces algorithmes aux caractéristiques de l'archive météorologique utilisée. Enfin, grâce à la mise en place d'une chaîne opérationnelle prototype, nous avons pu tester les performances opérationnelles de l'adaptation par analogie, notamment via l'étude de cas de l'évènement intense survenu les 8 et 9 septembre 2002 dans les Cévennes.

Mots clés :

prévision quantitative des précipitations ; prévision probabiliste ; adaptation des sorties des modèles météorologiques ; réanalyses NCEP/NCAR ; critère CRPS.